

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC973 U.S. PTO
09/840939
04/24/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日

Date of Application:

2000年10月 6日

出 願 番 号

Application Number:

特願2000-307907

願 人

Applicant (s):

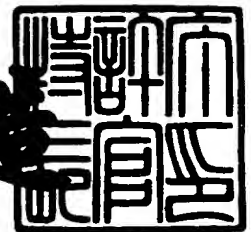
ソニー株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 3月16日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3020046

【書類名】 特許願

【整理番号】 0000776402

【提出日】 平成12年10月 6日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 7/24

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社
内

【氏名】 安部 素嗣

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社
内

【氏名】 西口 正之

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代表者】 出井 伸之

【代理人】

【識別番号】 100082131

【弁理士】

【氏名又は名称】 稲本 義雄

【電話番号】 03-3369-6479

【先の出願に基づく優先権主張】

【出願番号】 特願2000-127658

【出願日】 平成12年 4月24日

【手数料の表示】

【予納台帳番号】 032089

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9708842

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 信号処理装置及び方法

【特許請求の範囲】

【請求項 1】 少なくとも第 1 の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第 1 の信号の候補区間を検出する候補区間検出手段と、

上記候補区間内又はその前後の入力信号から、上記第 1 の信号らしさを表わす特徴量を抽出する特徴量抽出手段と、

上記特徴量に基づき、上記第 1 の信号の区間を検出する検出手段とを有することを特徴とする信号処理装置。

【請求項 2】 上記検出手段は、上記特徴量に基づいて上記候補区間が上記第 1 の信号である可能性を評価する特徴量評価手段と、上記評価結果に基づいて上記第 1 の信号の区間を判定する判定手段とを有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 3】 上記検出手段は、上記特徴量に基づいて、上記候補区間の信号と、別途指定した第 1 の信号との一致を判定する一致判定手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 4】 上記入力信号の振幅を検出する振幅検出手段を有し、
上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の振幅が所定の値より小さくなるパターンを検出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 5】 上記入力信号の変化を検出する変化検出手段を有し、
上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号が急激に所定の変化量を超えて変化するパターンを検出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 6】 上記入力信号の所定の信号成分が、所定の範囲内に収まる単位区間を検出する一様成分検出手段を有し、

上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の単位区間についての所定の信号成分が一様になるパターンを検

出することを特徴とする請求項1記載の信号処理装置。

【請求項7】 上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間の前及び／又は後の入力信号の振幅の大きさを抽出することを特徴とする請求項1記載の信号処理装置。

【請求項8】 上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間の前及び／又は後の入力信号の振幅が、所定の閾値より小さい区間の時間長を抽出することを特徴とする請求項1記載の信号処理装置。

【請求項9】 上記特徴量抽出手段は、上記入力信号の相関を検出する相関検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間における入力信号の相互相関を抽出することを特徴とする請求項1記載の信号処理装置。

【請求項10】 上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間における入力信号の平均振幅を抽出することを特徴とする請求項1記載の信号処理装置。

【請求項11】 上記特徴量抽出手段は、上記入力信号の変化を検出する変化検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間において入力信号が急激に変化する回数又は頻度を抽出することを特徴とする請求項1記載の信号処理装置。

【請求項12】 上記特徴量抽出手段は、上記入力信号の所定の信号成分が一様となる単位区間を検出する一様成分検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間において上記入力信号の所定の信号成分が一様となる単位区間の発生回数又は頻度を抽出することを特徴とする請求項1記載の信号処理装置。

【請求項13】 上記特徴量抽出手段は、複数のモードを取り得る入力信号の当該モードを検出するモード検出手段を有し、上記第1の信号らしさを表わす特徴量として、上記第1の信号の候補区間における上記モードを抽出することを特

徴とする請求項 1 記載の信号処理装置。

【請求項 1 4】 上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後に続く第 1 の信号の有無を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 1 5】 上記特徴量抽出手段は、上記入力信号のスペクトルを検出するスペクトル検出手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後の境界における上記スペクトルの変化を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 1 6】 上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、複数のチャンネルの何れかを取り得る入力信号の上記チャンネル情報を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 1 7】 上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、地域毎のコードを取り得る入力信号の当該地域コードを抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 1 8】 上記特徴量抽出手段は、上記入力信号の信号源を識別する信号源識別手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間における信号源の種類を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 1 9】 時間を計測する時計を有し、
上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間が入力される時刻を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 0】 上記特徴量抽出手段は、複数のジャンルに分け得る上記入力信号のジャンルを識別するジャンル識別手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前後の信号のジャンルを抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 1】 上記特徴量抽出手段は、時間を計測する時計と、複数のジャンルに分け得る上記入力信号の上記ジャンルを識別するジャンル識別手段とを有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の

前後の信号のジャンル、及び、上記第 1 の信号の候補区間の入力時刻からの経過時間を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 2】 上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、上記入力信号の振幅が基準値より小さいときの回数、その長さ、または上記入力信号の振幅の分散を抽出することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 3】 上記特徴量評価手段は、上記特徴量を荷重加算し、当該加重加算後の特徴量に基づき、上記候補区間が上記第 1 の信号である可能性を評価することを特徴とする請求項 2 記載の信号処理装置。

【請求項 2 4】 上記特徴量評価手段は、上記特徴量の評価の際に、多層パーセプトロンを用いることを特徴とする請求項 2 記載の信号処理装置。

【請求項 2 5】 上記入力信号を記録及び／又は再生する記録及び／又は再生手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 6】 上記入力信号を編集する編集手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 7】 上記第 1 の信号の区間をスキップするスキップ手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 8】 上記第 1 の信号の区間のみを取り出す取り出し手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 2 9】 上記入力信号は映像及び／又は音声信号からなり、上記第 1 の信号の区間はコマーシャルメッセージ区間であることを特徴とする請求項 1 記載の信号処理装置。

【請求項 3 0】 少なくとも第 1 の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第 1 の信号の候補区間を検出し、

上記候補区間内又はその前後の入力信号から、上記第 1 の信号らしさを表わす特徴量を抽出し、

上記特徴量に基づき、上記第 1 の信号の区間を検出することを特徴とする信号処理方法。

【請求項 3 1】 上記第 1 の信号区間の検出の際には、上記特徴量に基づいて上記候補区間が上記第 1 の信号である可能性を評価し、上記評価結果に基づいて上記第 1 の信号の区間を判定することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 2】 上記第 1 の信号区間の検出の際には、上記特徴量に基づいて、上記候補区間の信号と、別途指定した第 1 の信号との一致を判定することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 3】 上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の振幅が所定の値より小さくなるパターンを検出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 4】 上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号が所定の変化量を超えて急激に変化するパターンを検出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 5】 上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の単位区間についての所定の信号成分が所定の範囲内に収まるパターンを検出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 6】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間の前及び／又は後の入力信号の振幅の大きさを抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 7】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間の前及び／又は後の入力信号の振幅が、所定の閾値より小さい区間の時間長を抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 8】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における入力信号の相互相関を抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 3 9】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における入力信号の平均振幅を抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 40】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間において入力信号が急激に変化する回数又は頻度を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 41】 上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間において上記入力信号の所定の信号成分が一様となる単位区間の発生回数又は頻度を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 42】 複数のモードを取り得る入力信号の当該モードを検出し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における上記モードを抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 43】 上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後に続く第 1 の信号の有無を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 44】 上記入力信号のスペクトルを検出し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後の境界における上記スペクトルの変化を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 45】 上記第 1 の信号らしさを表す特徴量として、複数のチャンネルの何れかを取り得る入力信号の上記チャンネル情報を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 46】 上記第 1 の信号らしさを表す特徴量として、地域毎のコードを取り得る入力信号の当該地域コードを抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 47】 上記入力信号の信号源を識別し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間における信号源の種類を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 48】 上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間が入力される時刻を抽出することを特徴とする請求項 30 記載の信号処理方法。

【請求項 49】 複数のジャンルに分け得る上記入力信号の上記ジャンルを識別し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の

前後の信号のジャンルを抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 5 0】 複数のジャンルに分け得る上記入力信号のジャンルを識別し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前後の信号のジャンル、及び、上記第 1 の信号の候補区間の入力時刻からの経過時間を抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 5 1】 上記第 1 の信号らしさを表す特徴量として、上記入力信号の振幅が基準値より小さいときの回数、その長さ、または上記入力信号の振幅の分散を抽出することを特徴とする請求項 3 0 記載の信号処理方法。

【請求項 5 2】 上記特徴量に基づく評価の際には、上記特徴量を荷重加算し、当該加重加算後の特徴量に基づき、上記候補区間が上記第 1 の信号である可能性を評価することを特徴とする請求項 3 1 記載の信号処理方法。

【請求項 5 3】 上記特徴量に基づく評価の際には、多層パーセプトロンを用いることを特徴とする請求項 3 1 記載の信号処理方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、信号処理装置及び方法に関し、特に、テレビジョン放送に付加されるコマーシャルメッセージを自動的に検出或いは検索可能とする信号処理装置及び方法に関するものである。

【0 0 0 2】

【従来の技術】

一般に、放送番組についてスポンサーが付いているテレビジョン（以下、適宜 TV とする）放送では、本編（番組そのもの）の合間に、その番組を提供しているスポンサーのコマーシャルメッセージ（以下、単に CM とする。）が挿入される。

【0 0 0 3】

しかしながら、視聴者の興味は、主として本編にあるため、CM を省略して視聴したいという要望は多い。

【0004】

その一方で、関心のある製品のCMや、ストーリー性のあるCM、有名俳優が登場するCMなど、CM自体を視聴対象とする場合も少なくない。

【0005】

このようなことから、特に、磁気テープや磁気ディスク等の媒体にTV放送番組等を録画する場合において、放送番組から本編とCMとを分離／区別して、必要なときに必要な部分を視聴することが可能となれば、上述したような多様な要望に対する有効な解決手段を与えることになると考えられる。

【0006】

ところで、TV放送からCMのみを省略して試聴可能とする技術としては、従来より、いわゆるCMスキップ機能（広義のCMスキップ機能）が存在し、例えば家庭用ビデオ録画再生装置には当該CMスキップ機能が搭載されているものが多い。

【0007】

このCMスキップ機能（広義のCMスキップ機能）は、その仕組みに応じて4つに大別され、それぞれ一般には、CM早送り機能、CMカット機能、放送モードに基づくCMスキップ機能（狭義のCMスキップ機能）、放送モードによらないCMスキップ機能（狭義のCMスキップ機能）と呼ばれている。なお、これら4つに大別されたCMスキップ機能のうち、上記CM早送り機能とは、日本国内においては通例としてCMが15秒の倍数で製作されていることを利用し、家庭用ビデオ録画再生装置でのビデオ再生中に、例えばリモコンボタン操作によって（30秒）ラ（ボタンを押した回数）の時間だけ早送りを行うことにより、CM部分を飛ばす機能である。このCM早送り機能に関する技術については、例えば特開平10-269651号公報などに公開されている。また、特開平9-307841号公報には、上記CM早送り機能における早送り終了時刻を時間で決定せず、映像信号のブラックフレームと音声信号の無音部が同時に所定長さ以上にわたり生じる部分を、早送り終了時刻とする技術が公開されている。

【0008】

CMカット機能とは、日本国内においてはCMがステレオ放送されることが多

く、また、TV放送信号には音声モノラルモード／ステレオモード／音声多重モードの何れであるかを示すパイロット信号が多重されていることを利用し、例えば音声モノラルモード又は音声多重モードとなされている本編を録画するような場合に、ステレオモード区間（すなわちCM区間）のみ録画を停止することで、CM部分をカットすることを可能にする機能である。このCMカット機能に関する技術については、例えば特開平3-158086号、特開平3-2622872号などの公報に公開されている。

【0009】

放送モードに基づくCMスキップ機能（狭義のCMスキップ機能）とは、CMカット機能と同様に放送モードの違いを利用する仕組みであるが、上記CMカット機能が録画時にCMをカットしているのに対し、当該放送モードに基づくCMスキップ機能では、録画時には全ての映像及び音声を記録すると同時にその放送モードも記録しておき、再生時にステレオ放送区間（すなわちCM区間）のみ自動的に早送りする機能である。この放送モードに基づくCMスキップ機能に関する技術については、例えば特開平5-250762号公報などに公開されている。

【0010】

放送モードによらないCMスキップ機能（狭義のCMスキップ機能）とは、放送信号中に含まれる音声信号の無音区間間隔や、映像信号の映像変換点（映像が急激に変化する点）の出現間隔、映像信号の黒レベル／白レベルの出現間隔などを利用して、これらが15秒の倍数間隔に合致する部分をCMとして早送りする機能である。当該放送モードによらないCMスキップ機能に関する技術については、特開平8-317342号公報や、文献「TV放送のCM検出方式についての一検討」（映像情報メディア学会技術報告、VIR97-22、19/23（1997））などに公開されている。

【0011】

【発明が解決しようとする課題】

しかしながら、上記CM早送り機能自体は、家庭用ビデオ録画再生装置にて録画したTV放送を再生して視聴する視聴者に対して、CMかどうかの判断を任せ

るものであり、TV放送から自動的にCM部分を検出するようなCM検出機能は備えていない。

【0012】

また、上記放送モードに基づくCMカット機能やCMスキップ機能（狭義のCMスキップ機能）では、放送モードの違い、すなわち例えばステレオモードによるCM検出を行うようになされているが、例えば本編自体がステレオモードとなされている場合や、CMがモノラルモードや音声多重モードとなっている場合には、全く効果がない（すなわちCMカットやCMスキップができない）。

【0013】

一方、放送モードによらないCMスキップ機能（狭義のCMスキップ機能）では、視聴者や放送モードに依存することなく、放送内容に基づいたCM検出が可能となっている。ただし、当該機能の場合は、音声信号の無音区間間隔や、映像信号の映像変換点（映像が急激に変化する点）の出現間隔などが、予め想定されている設定値と一致することなど、非常に狭い条件に基づいた決定論的な判定手続きによりCM区間を検出するようになされている。これに対し、実際の放送番組では、放送時間の調整の目的や人為的なスイッチングなどの要因によって無音区間間隔が例えば短縮されていたりする場合が少なからずあり、逆に、番組本編の中にも上記の条件を満たす区間が多く存在している。このため、無音区間間隔が予め想定されている長さより短いようなCMの場合は全く検出ができなくなる問題が発生し、逆に、上記設定値の条件を満たす本編の場合は当該本編をCM区間として誤って検出してしまう等の問題点が生じてしまう。

【0014】

さらに、上記4つの機能では、例えば複数のCMが連続されて放送されている場合には、CM全体の区間を検出することはできても、個々のCMの検出はできない。このため、CMを個別に抽出して視聴したいような場合（要望）には適さない。

【0015】

以上のようなことから、TV放送信号からCM部分を高精度に検出又は検索できることが望まれている。

【 0 0 1 6 】

そこで、本発明は、以上のような状況を鑑みてなされたものであり、例えばTV放送信号に含まれるコマーシャルメッセージ部分を高精度に検出又は検索可能とする、信号処理装置及び方法を提供することを目的とする。

【 0 0 1 7 】

【課題を解決するための手段】

本発明の信号処理装置は、少なくとも第1の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて第1の信号の候補区間を検出する候補区間検出手段と、その候補区間内又はその前後の入力信号から第1の信号らしさを表わす特徴量を抽出する特徴量抽出手段と、その特徴量に基づき第1の信号の区間を検出する検出手段とを有することにより、上述した課題を解決する。

【 0 0 1 8 】

ここで、検出手段は、特徴量に基づいて候補区間が第1の信号である可能性を評価する特徴量評価手段と、評価結果に基づいて第1の信号の区間を判定する判定手段とを有する。或いは、検出手段は、特徴量に基づいて、候補区間の信号と、別途指定した第1の信号との一致を判定する一致判定手段を有する。

【 0 0 1 9 】

次に、本発明の信号処理方法は、少なくとも第1の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第1の信号の候補区間を検出し、その候補区間内又はその前後の入力信号から第1の信号らしさを表わす特徴量を抽出し、その特徴量に基づき第1の信号の区間を検出することにより、上述した課題を解決する。

【 0 0 2 0 】

ここで、第1の信号区間の検出の際には、特徴量に基づいて候補区間が第1の信号である可能性を評価し、評価結果に基づいて第1の信号の区間を判定する。或いは、第1の信号区間の検出の際には、特徴量に基づいて、候補区間の信号と、別途指定した第1の信号との一致を判定する。

【 0 0 2 1 】

すなわち本発明は、例えばTV放送信号の本編とコマーシャルメッセージ部分のように、本編信号とコマーシャルメッセージ信号（第1の信号）よりなるTV放送信号の音声及び／又は映像信号から、所定の必須条件に基づき、信号が特徴的パターンを示すCM候補区間を抽出し、それらCM候補区間に対して、幾つの特徴からCMらしさを表わす指標である付加条件を検出し、抽出した付加条件を総合してCMの判定を行うことにより、精度良く個々のCMの検出を可能とするものである。また、本発明は、上記検出する付加条件の一部又は全部を記録し、TV放送信号や記録された他のCM候補と比較することより、TV放送信号又はそれが記録されたデータベースから所望のCMを検索可能とするものである。

【0022】

【発明の実施の形態】

以下、本発明の好ましい実施の形態について、図面を参照しながら説明する。

【0023】

先ず、本発明実施の形態の具体的構成について述べる前に、本発明に基づき、例えばTV放送信号からCM部分を検出する際の原理について概説する。

【0024】

一般に、TV放送されるCMは、放送者の指定する規格に基づいて製作されるため、その「時間長（1つのCMの時間）はごく少数の種類に限定」される。例えば日本国内においては、特殊な例を除くほぼ全てのCMが、15秒、30秒、60秒の長さで製作されている。

【0025】

また、CMの放送時には、番組本編や他のCMとは独立に製作されたものがTV放送ストリーム中に挿入されるため、各CMの前後では必然的に「音声レベルが下がる（すなわち小音量となる）」こと、及び、「映像信号が切り替わる」こと、という特徴を持つ。ここで、「音声レベルが下がる」とは、必ずしも無音（ここでは極微小なノイズしかない部分という意味）と同義ではない。すなわち、実際には、CMと本編との切り替えのタイミングなどにより、必ずしも完全に無音とはならないまま切り替わることがあるからである。

【0026】

上述したように、CMについての「規定時間長（少数種類の時間長）」、「小音量」、「映像切り替わり」という3つの特徴は、ほぼ全てのCMが示すパターンである。本発明では、このようにほぼ全てのCMが示す信号のパターンを「特徴的パターン」と称し、それを検出するための条件を「必須条件」と称することにする。

【0027】

したがって、TV放送信号から当該必須条件に対応する信号部分を検出するようにすれば、誤棄却がほとんどなく決定論的にCMの候補（すなわちCMであろうと思われる信号部分）を検出することが可能となる。但し、番組本編内にも、偶然そのような必須条件を満たしてしまう部分が多く存在するため、上記必須条件を用いただけでは、番組本編の一部をCM候補として誤検出してしまう虞が残る。

【0028】

一方で、上記必須条件と比べて例外は多く存在するものの、CMの性質上、多くのCMが満たすか若しくは一定の傾向を示す特徴としては、以下のようなものがある。

【0029】

1) CMの前後（CMが開始される直前とCMが終了して本編番組が開始又は再開される直前）では、通常の番組本編内よりも音声レベルが低くなることが多い。

【0030】

2) CMと番組本編との間、及び、あるCMと他のCMとの間の、ほぼ無音となる区間長は、数百ミリ秒程度であることが多い。

【0031】

3) TV放送内に含まれる有音区間は、CMの規定時間長（15秒、30秒、60秒等）より百ミリ秒程度以上短いことが多く、また1秒程度以上短いことは少ない。

【0032】

4) ステレオ音声信号の左チャンネル（Lチャンネル）と右チャンネル（Rチャ

ネル)の相関値は、1より有意に小さいことが多い。

【0033】

5) CM期間中は、番組本編より音量が大きめである傾向がある。

【0034】

6) CMの放送モードは、ステレオモードであることが多い。

【0035】

7) CM区間では、複数のCMが連続して放送されることが多い。

【0036】

8) CM期間中は、映像カットの切り替わり頻度が高いことが多い。

【0037】

9) 逆に、CM期間中であっても、カットの切り替わり頻度が極端に低いものがある(例えば静止画によるCMなど)。

【0038】

10) 番組本編とCMとの境界や、あるCMと他のCMの境界では、音質が大きく変化することが多い。

【0039】

11) CMは、音声と音楽を同時に含むことが多い。

【0040】

12) 番組編成上、毎時丁度の時刻近辺では、CMが放送される確率が高い。

【0041】

13) 同様に、毎時30分付近でもCMが放送される確率が高い。

【0042】

14) 番組のジャンルによってCMが放送される確率の高い時間帯がある(例えばサッカー中継のハーフタイムなど)。

【0043】

本発明では、以下、これらの特徴に基づく条件を「付加条件」と称することにする。すなわち、当該付加条件は、CMが、規格に基づいて製作されるという制約、短い時間で宣伝効果を上げるためのものであるという制約、及び、番組構成上の都合などによる制約の元で製作された結果として、TV放送信号上に現れて

くることによる条件である。したがって、この付加条件は、決定論的な取り扱いができるほど確実な条件ではないものの、CMである可能性（CMらしさ）を評価する際の有効な条件となる。

【 0 0 4 4 】

さらに、TV放送においては、同時に同じチャンネルで複数の映像及び音声が発送されることは物理的にありえないという特徴がある。すなわち、TV放送信号からCMであろうと思われる信号部分（CM候補）を検出しようとする場合において、例えば、TV放送信号中に、上記付加条件を満たす複数の映像及び音声区間がオーバーラップして存在し、何らかの処理の結果、当該オーバーラップ区間でCM候補が検出されたとしても、そのオーバーラップしている複数の映像及び音声内の少なくともどちらかの区間は、正しいCM区間ではあり得ない。本発明では、TV放送におけるこのような特徴に基づく条件を、「論理条件」と称することにする。

【 0 0 4 5 】

本発明では、以上説明した「必須条件」、「論理条件」、「付加条件」を合理的かつ効果的に利用することにより、TV放送信号から高精度でCM部分を検出可能としている。

【 0 0 4 6 】

より具体的に言うと、本発明では、「必須条件」に基づき、決定論的にTV放送信号中からCM候補（CMであろうと思われる信号部分）を抽出し、「付加条件」に基づくCMらしさ（CMである可能性）の統計論的な評価によってCM候補を選択し、「論理条件」によりCM候補のオーバーラップ関係を解消することにより、精度の高いCM検出を実現するものである。

【 0 0 4 7 】

図1には、上述のようにTV放送信号からCMを検出可能とした、本発明の第1の実施の形態としての映像音声記録装置の概略構成を示す。

【 0 0 4 8 】

図1において、アンテナ6により受信されたTV放送信号は、チューナ1により同調され、放送信号1aとして復調器2に入力する。

【 0 0 4 9 】

復調器 2 は、入力された放送信号 1 a に含まれるパイロット信号の変調周波数から、その入力時点での放送モード（モノラルモード／ステレオモード／音声多重モード）を判別し、また、放送信号 1 a に含まれる映像及び音声の変調信号を復調する。このとき、音声変調信号は、上記判別した放送モードに応じて復調される。当該復調器 2 により復調された映像信号 2 a、音声信号 2 b、及び放送モード信号 2 c は、CM 検出部 4 に入力される。また、映像信号 2 a、音声信号 2 b は、映像音声記録部 5 にも入力される。

【 0 0 5 0 】

映像音声記録部 5 は、磁気テープ、磁気ディスク、光ディスク若しくは光磁気ディスク等を記録媒体として用い、上記映像信号 2 a 及び音声信号 2 b を記録再生可能な装置である。なお、映像音声記録部 5 は、映像信号及び音声信号に対して符号化や圧縮などを施して記録し、再生時にはその符号化や圧縮された映像信号及び音声信号を復号化、伸張する装置であってもよい。

【 0 0 5 1 】

時計部 3 は、時刻を表す時間信号 3 a を発生し、当該時間信号 3 a を CM 検出部 4 に供給する。

【 0 0 5 2 】

CM 検出部 4 は、本発明の主要部であり、詳細については後述するが、上記復調器 2 からの映像信号 2 a、音声信号 2 b、放送モード信号 2 c と、時計部 3 からの時間情報 3 a とを入力とし、放送信号 1 a の映像及び音声信号に含まれる CM 区間を検出し、その開始時刻と長さを CM 検出出力 4 a として出力する。当該 CM 検出部 4 による CM 検出出力 4 a は、映像音声記録部 5 に送られ、上記映像信号 2 a 及び音声信号 2 b と共に記録される。

【 0 0 5 3 】

また、この CM 検出部 4 には、チューナ 1 から、上記同調された放送チャンネルを示すチャンネル情報 1 b も供給される。当該チャンネル情報 1 b は、CM が含まれない放送チャンネルが指定されたか否かを、CM 検出部 4 において判断する為に用いられる。すなわち、CM 検出部 4 は、上記チャンネル情報 1 b により

、上記チューナ 1 において明らかに CM が放送されない放送チャンネルの指定がなされていると判断した場合、上記 CM 検出動作を行わないようにする。なお、ここでは、CM 検出部 4 自身がチャンネル情報 1 b に基づいて CM 検出動作を行うか否か判断することとしているが、チューナ 1 がチャンネル情報 1 b により CM 検出部 4 での CM 検出動作を制御するようにしても良い。

【 0 0 5 4 】

以上、チューナ 1、復調器 2、時計部 3、映像音声記録部 5 に示した各構成要素は、現在広く知られているビデオ装置などの磁気記録再生装置のものと同様であるため、以下では、CM 検出部 4 に限って詳細に説明する。

【 0 0 5 5 】

図 2 には、上記 CM 検出部 4 の第 1 の具体例の詳細な構成を示す。なお、この図 2 中の各信号のうち、図 1 と共通の信号については、図 1 と同じ指示符号を付している。また、当該 CM 検出部 4 は、大別して、フロントエンド部とバックエンド部とから構成されている。また、図中の動作制御部 2 3 は、チューナ 1 から供給された上記チャンネル情報 1 b に基づいて、上記チューナ 1 において明らかに CM が放送されない放送チャンネルの指定がなされているか否か判断し、その判断結果に応じて、当該図 2 の各部における CM 検出動作を行わないように制御するものである。

【 0 0 5 6 】

先ず、図 2 のフロントエンド部から説明する。

【 0 0 5 7 】

この図 2 において、図 1 の復調器 2 より供給された映像信号 2 a は、A/D 変換器 1 0 にてデジタル化され、フレームメモリ 1 1 に蓄えられる。なお、フレームメモリ 1 1 は、少なくとも 2 フレーム分の映像信号を蓄積可能なメモリである。当該フレームメモリ 1 1 からフレーム毎に読み出された映像信号は、カットチェンジ検出器 1 2 に送られる。

【 0 0 5 8 】

カットチェンジ検出器 1 2 は、フレームメモリ 1 1 より供給されたフレーム毎の映像信号に基づいて、映像が急激に変化するフレーム（以下、映像変化フレー

ムと呼ぶ)と、輝度が一樣となるフレーム(以下、一樣輝度フレームと呼ぶ)を検出する。

【0059】

すなわち、カットチェンジ検出器12は、フレームメモリ11に蓄えられた時間的に隣接する2つのフレーム映像間で、各画素毎に輝度の差分の自乗和を求め、当該自乗和が所定の閾値を越えた場合に、上記隣接する2つのフレームのうちの時間的に後のフレームを、上記映像が急激に変化する映像変化フレームとして検出する。また、カットチェンジ検出器12は、フレームメモリ11に蓄えられた各フレーム映像の輝度の分散を求め、その輝度の分散値が所定の閾値以下である場合に、そのフレームを一樣輝度フレームであるとして検出する。なお、フレームの間隔(NTSC方式では約30ms)が、後述する音声信号処理において説明するフレーム周期と一致しない場合には、当該フレーム間隔を再離散化することによって、フレーム周期と一致させておくようにする。

【0060】

以下、当該カットチェンジ検出器12における映像変化フレームと一樣輝度フレームの検出について、より具体的に説明する。

【0061】

ここで、離散化された映像信号の横サイズを X 、縦サイズを Y 、縦横の画素番号を x, y とし、第 n フレームの映像を $I_n(x, y)$ 、当該第 n フレームに対して時間的に1フレーム前の第 $n-1$ フレームの映像を $I_{n-1}(x, y)$ として表わすと、第 n フレームと第 $n-1$ フレームの間の各画素毎の輝度差分の自乗和 $D[n]$ は、式(1)により得られ、また、第 n フレームの輝度分散値 $V[n]$ は、式(2)により得られる。

【0062】

【数 1】

$$D[n] = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left(I_n(x,y) - I_{n-1}(x,y) \right)^2 \quad (1)$$

$$V[n] = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I_n^2(x,y) - \left(\frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I_n(x,y) \right)^2 \quad (2)$$

【0063】

また、このときのカットチェンジ検出器 12 の検出出力 $C[n]$ は、式 (3) により表わされる。

【0064】

【数 2】

$$C[n] = \begin{cases} 1 & (D[n] \geq D_{thsd} \text{ or } V[n] \leq V_{thsd}) \\ 0 & (D[n] < D_{thsd} \text{ and } V[n] > V_{thsd}) \end{cases} \quad (3)$$

【0065】

ただし、式中の D_{thsd} は上記映像変化フレームを検出する際の前記自乗和に対する所定の閾値であり、 V_{thsd} は上記一様輝度フレームを検出する際の前記輝度の分散値に対する所定の閾値である。

【0066】

当該カットチェンジ検出器 12 の検出出力 $C[n]$ は、映像信号についての特徴量として特徴量バッファ 18 へ送られる。

【0067】

なお、上記の 2 つのフレーム映像間で輝度差分を求める際には、2 フレーム分の映像信号を蓄積可能なメモリが必要となり、また、2 フレーム分の映像信号に対する演算量も必要となる。そこで、例えばフレーム映像全面を同時に処理する代わりに、フレーム映像を適切な小ブロック毎に分け、その小ブロック毎に輝度差分を求めるようにしたり、或いは、フレーム映像間の各画素毎に輝度差分を求めるのではなく、各フレーム映像毎に輝度ヒストグラムを求めて、その輝度ヒス

トグラムフレーム間差分を求めるようにしたり、又は、各フレーム映像毎に平均輝度を求めて、その平均輝度のフレーム間差分を求めるようにするで、メモリ容量や演算量を減らすことも可能である。逆に、メモリや演算量に余裕がある場合には、例えば、カラー映像におけるR（赤）、G（緑）、B（青）成分のようなカラー成分毎に、上記輝度差分やカラーヒストグラム差分を求めることで、より検出精度を高めることも可能である。

【 0 0 6 8 】

次に、図1の復調器2より供給された音声信号2bは、A/D変換器13にてデジタル化され、音声信号バッファ14に蓄えられる。なお、音声信号バッファ14は、少なくとも所定時間 T_1 （例えば30ms、以下、これを1フレーム長とする）分の左（L）右（R）2チャンネルのステレオ音声信号を蓄積可能なメモリである。当該音声信号バッファ14から読み出された音声信号は、振幅検出器15、相関検出器16、スペクトル検出器17に送られる。

【 0 0 6 9 】

振幅検出器15は、音声信号バッファ14に蓄えられた音声信号を用いて、所定の時間 T_2 （例えば15ms、以下、これを1フレーム周期とする）毎の短時間平均自乗振幅を検出する。すなわち、振幅検出器15は、音声信号バッファ14に左右2チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ14より読み出された左右2チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、所定の時間 T_2 （15ms、1フレーム周期）毎に、短時間平均自乗振幅を検出する。なお、上記 m （ $m=0, \dots, M-1$ ）は、離散化された時間を表わすバッファ内のサンプル番号であり、最大番号 M が1フレーム長 T_1 に対応する。

【 0 0 7 0 】

より具体的に説明すると、振幅検出器15は、第 n フレームにおける左右2チャンネルの音声信号の平均自乗振幅 $A[n]$ を式（4）により計算する。すなわち、平均自乗振幅は15ms（1/2フレーム）毎に計算され、その15ms毎の平均自乗振幅の30ms（1フレーム）の期間における平均値がさらに演算され、最終的な、1フレームの平均自乗振幅とされる。

【 0 0 7 1 】

【数 3】

$$A[n] = \frac{1}{4M} \sum_{m=0}^{M-1} \left(S_L[m+nT_2] + S_R[m+nT_2] \right)^2 \quad (4)$$

【 0 0 7 2 】

当該振幅検出器 1 5 の検出出力である平均自乗振幅 $A[n]$ は、音声信号についての特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 0 7 3 】

相関検出器 1 6 は、音声信号バッファ 1 4 に蓄えられた音声信号を用いて、1 フレーム毎の音声信号について規格化前の相関係数を検出すると共に、後段にて行われる規格化のための短時間エネルギーも同時に検出する。すなわち、相関検出器 1 6 は、音声信号バッファ 1 4 に左右 2 チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ 1 4 より読み出された左右 2 チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、1 フレーム毎の左右 2 チャンネルの音声信号について規格化前の相関係数を検出すると共に、後段にて行われる規格化のための短時間エネルギーも同時に検出する。

【 0 0 7 4 】

より具体的に説明すると、相関検出器 1 6 は、第 n フレームにおける左右 2 チャンネルの音声信号の相関係数 $A_{LR}[n]$ を式 (5) により計算し、左チャンネルの音声信号エネルギー $A_{LL}[n]$ を式 (6) により計算し、右チャンネルの音声信号エネルギー $A_{RR}[n]$ を式 (7) により計算する。

【 0 0 7 5 】

【数 4】

$$A_{LR}[n] = \sum_{m=0}^{M-1} S_L[m+nT_2] S_R[m+nT_2] \quad (5)$$

$$A_{LL}[n] = \sum_{m=0}^{M-1} S_L^2[m+nT_2] \quad (6)$$

$$A_{RR}[n] = \sum_{m=0}^{M-1} S_R^2[m+nT_2] \quad (7)$$

【 0 0 7 6 】

当該相関検出器 1 6 の検出出力である相関係数 $A_{LR}[n]$ と音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ は、それぞれが音声信号についての特徴量のひとつとして特徴量バッファ 1 8 へ送られる。

【 0 0 7 7 】

スペクトル検出器 1 7 は、音声信号バッファ 1 4 に蓄えられた音声信号を用いて、短時間スペクトルを計算する。すなわち、スペクトル検出器 1 7 は、音声信号バッファ 1 4 に左右 2 チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ 1 4 より読み出された左右 2 チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、短時間スペクトルを計算する。

【 0 0 7 8 】

より具体的に説明すると、スペクトル検出器 1 7 は、第 n フレームにおける左右 2 チャンネルの音声信号の離散スペクトル $F[k;n]$ を求める。なお、 $k=0$ 、 \dots 、 $K-1$ を離散化された周波数を表わす番号とすると、離散スペクトル $F[k;n]$ は式 (8) により表わされる。

【 0 0 7 9 】

【数 5】

$$F[k;n] = \left| \sum_{m=0}^{M-1} (S_L[m] + S_R[m]) e^{-2\pi j m k / M} \right|^2 \quad (8)$$

【 0 0 8 0 】

この式 (8) の演算は、例えば、高速フーリエ変換 (FFT) 又は線形予測分析 (LPC) などを用いて実現される。

【 0 0 8 1 】

当該スペクトル検出器 1 7 の計算出力である短時間離散スペクトル $F[k;n]$ は、音声信号についての特徴量のひとつとして特徴量バッファ 1 8 へ送られる。

【 0 0 8 2 】

次に、図 1 の復調器 2 より供給された放送モード信号 2 c は、上述した音声信号処理のフレームに合わせて離散化された数値となされる。

【 0 0 8 3 】

より具体的に説明すると、第 n フレームにおける放送モード信号 2 c は、例えば式 (9) のような数値 $B[n]$ となされる。

【 0 0 8 4 】

【数 6】

$$B[n] = \begin{cases} 0 & (\text{モノラルモード}) \\ 1 & (\text{ステレオモード}) \\ 2 & (\text{音声多重モード}) \end{cases} \quad (9)$$

【 0 0 8 5 】

この放送モード信号 2 c を離散化した数値 $B[n]$ は、TV 放送信号の特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 0 8 6 】

同様に、図 1 の時計部 3 より供給された時間信号 3 a も、音声信号処理のフレームに合わせて離散化された数値 $T[n]$ となされ、特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 0 8 7 】

特徴量バッファ 1 8 は、上記カットチェンジ検出器 1 2 からの検出出力 $C[n]$ と、振幅検出器 1 5 からの平均自乗振幅 $A[n]$ と、相関検出器 1 6 からの相関係数 $A_{LR}[n]$ 、音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ と、スペクトル検出器 1 7 からの短時間離散スペクトル $F[k;n]$ と、放送モード信号 2 c の離散化数値 $B[n]$ と、時間信号 3 a の離散化数値 $T[n]$ とからなる、式 (1 0) に示される特徴量 $G[n]$ を、所定の時間 T_3 に渡って蓄積する。なお、時間 T_3 は CM 部分を最低でも 1 つ以上に渡って記憶できる時間であり、例えば 8 0 秒などとする。

$$G[n] = \{C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n], B[n], T[n]\} \quad (10)$$

【 0 0 8 8 】

以上の A/D 変換器 1 0 から特徴量バッファ 1 8 までが、図 2 に示した CM 検出部 4 のフロントエンド部の構成であり、以下、図 3、図 4 のフローチャートを用いて当該フロントエンド部における処理の流れを説明する。なお、図 3 のステップ S 3 0 乃至 S 3 2 までは映像信号 2 a についての処理の流れを表しており、

図4のステップS33乃至S40までは音声信号2b及び放送モード信号2c、時間信号3aについての処理の流れを表している。

【0089】

先ず、映像信号2aについての処理の流れを表す図3において、フロントエンド部は、ステップS30の処理として、A/D変換器10によりデジタル化された、少なくとも1フレーム分の映像信号2aをフレームメモリ11に蓄える。このフレームメモリ11は、1フレーム分の映像信号2aを1サンプルとして扱うようになされており、1フレーム分の映像信号2aが入力されると、当該フレームメモリ11内に既に蓄積されている映像信号2aが1フレーム分シフトし、最も時間的に過去に入力された1フレームの映像信号2aが押し出されて出力されるようになっている。

【0090】

次に、フロントエンド部は、ステップS31の処理として、フレームメモリ11から映像信号2aを読み出してカッチェンジ検出器12に送り、前述のようにして検出出力C[n]を求める。

【0091】

その後、フロントエンド部は、ステップS32の処理として、当該検出出力C[n]を特微量バッファ18に蓄える。

【0092】

一方、音声信号2bについての処理の流れを表す図4において、フロントエンド部は、ステップS33及びステップS34の処理として、A/D変換器13によりデジタル化された、音声信号2bを音声信号バッファ14に入力すると共に、当該音声信号バッファ14に少なくとも1フレーム周期 T_2 分の音声信号2bを蓄積する。この音声バッファ14は、1フレーム周期 T_2 分の音声信号2bを1サンプルとして扱うようになされており、1フレーム周期 T_2 分の音声信号2bが入力されると、当該音声バッファ14内に既に蓄積されている音声信号2bが1フレーム周期 T_2 分だけシフトし、最も時間的に過去に入力された1フレーム周期 T_2 分の音声信号2bが押し出されて出力されるようになっている。

【0093】

上記音声信号バッファ 1 4 に少なくとも 1 フレーム周期 T_2 分の音声信号 2 b が蓄積されると、フロントエンド部は、ステップ S 3 5 の処理として、当該音声信号バッファ 1 4 に蓄積された音声信号 2 b を読み出して振幅検出器 1 5 に送り、前述のようにして、平均自乗振幅 $A[n]$ を求める。

【 0 0 9 4 】

同時に、フロントエンド部は、ステップ S 3 6 の処理として、音声信号バッファ 1 4 に蓄積された音声信号 2 b を相関検出器 1 6 に送り、前述のようにして、相関係数 $A_{LR}[n]$ と音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ を求める。

【 0 0 9 5 】

また同時に、フロントエンド部は、ステップ S 3 7 の処理として、音声信号バッファ 1 4 に蓄積された音声信号 2 b をスペクトル検出器 1 7 に送り、前述のようにして、短時間離散スペクトル $F[k;n]$ を求める。

【 0 0 9 6 】

さらに、フロントエンド部は、ステップ S 3 8 の処理として、図 1 の復調器 2 より供給された放送モード信号 2 c から、前述のように離散化した数値 $B[n]$ を求めると共に、図 1 の時計部 3 より供給された時間信号 3 a から、前述のように離散化された数値 $T[n]$ を求める。

【 0 0 9 7 】

フロントエンド部は、以上のようにして求められた、上記カットチェンジ検出器 1 2 からの検出出力 $C[n]$ と、振幅検出器 1 5 からの平均自乗振幅 $A[n]$ と、相関検出器 1 6 からの相関係数 $A_{LR}[n]$ 、音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ と、スペクトル検出器 1 7 からの短時間離散スペクトル $F[k;n]$ と、放送モード信号 2 c の離散化数値 $B[n]$ と、時間信号 3 a の離散化数値 $T[n]$ とからなる特徴量 $G[n]$ を、特徴量バッファ 1 8 に蓄積する。

【 0 0 9 8 】

図 2 に戻り、バックエンド部の説明を行う。なお、以下の説明において、番号 n は、特徴量バッファ 1 8 内にフレーム毎に蓄積される特徴量の、各フレーム番号を表わすものとする。また、最新のフレームの特徴量を $G[0]$ とし、過去のフレームの特徴量となるにしたがって n の値が増加し、新たなフレームの特徴量が

入力された場合には、全てのデータが1ずつシフト（フレーム番号が1ずつシフト）するものとする。

【0099】

図2において、特徴量バッファ18に蓄積された特徴量は、フレーム毎にCM候補検出器19に送られる。

【0100】

当該CM候補検出器19は、ほぼ全てのCMが満たす、前述した「必須条件」に基づき、フレーム毎にCM区間の候補を算出する。ここで、必須条件とは、前述したように、CMの音声信号が「小音量」であること、すなわち音声信号の音量が所定の閾値以下となっているフレーム（以下、音量条件と呼ぶ）であり、且つ、CMの「映像切り替わり」があること、すなわち映像信号が急激に変換するフレーム又は一様な輝度となるフレーム（以下、映像条件と呼ぶ）であり、さらに、「規定時間長（少数種類の時間長）」であること、すなわち上記音量条件と映像条件を満たす2つのフレームの間隔が所定のCM長と合致する区間（以下、時間条件と呼ぶ）となるような条件であり、具体的には、前述の特徴量を用いて、以下のような式（11）で且つ式（12）で且つ式（13）の条件として書き下すことができる。

【0101】

$$A[0] < A_{thsd} \quad (11)$$

$$C[0] = 1 \quad (12)$$

$$A[n_1] < A_{thsd}, C[n_1] = 1 \text{ 又は } A[n_2] < A_{thsd}, C[n_2] = 1 \text{ 又は } A[n_3] < A_{thsd}, C[n_3] = 1 \quad (13)$$

ただし、 A_{thsd} は所定の自乗振幅の閾値であり、 n_1 、 n_2 、 n_3 はそれぞれCM長として規定されている時間長（本実施の形態では、一例として15秒、30秒、60秒の3種類の時間長がある場合を説明に用いている）を、フレーム周期単位に換算した数である。なお、CMの実際の放送時間には誤差があるため、実用上は、 n_1 、 n_2 、 n_3 にはそれぞれ多少の幅を持たせる。

【0102】

ここで、図5を用いて、上記CM候補検出器19の動作の流れを説明する。

【 0 1 0 3 】

図5において、特徴量バッファ18では、ステップS50のバッファシフト処理とステップS51の特徴量入力処理として、図3のステップS32で説明したフレームメモリと図4のステップS40で説明した音声信号バッファと同様に、1フレーム単位の入力、シフト及び出力の動作を行うようになされている。すなわち、特徴量バッファ18は、1フレーム分の特徴量を1サンプルとして扱うようになされており、1フレーム分の特徴量が入力されると、当該特徴量バッファ18内に既に蓄積されている特徴量が1フレーム分だけシフトし、最も時間的に過去に入力された1フレーム分の特徴量が押し出されて出力されるようになっている。

【 0 1 0 4 】

上記ステップS50及びステップS51の処理により、特徴量バッファ18から1フレーム（1サンプル）分の特徴量が入力されると、CM候補検出器19は、ステップS52及びステップS53の処理として、1フレーム（サンプル）に特徴量が上記必須条件の音量条件、映像条件、時間条件を満たすか否かの評価を行う。すなわち、CM候補検出器19は、ステップS52において、先ず最初のフレームの平均自乗振幅 $A[0]$ と所定の自乗振幅の閾値 A_{thsd} を比較し、次に、ステップS53の処理として、前記検出出力 $C[0]$ が1となるか否か調べることにより、当該フレームが上記必須条件である音量条件、映像条件、時間条件を満たすか否かの判定を行う。CM候補検出器19では、これらステップS52、S53の判定処理の結果、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 A_{thsd} を超えず、且つ、上記必須条件を満たしていると判定した場合、当該フレームをCM候補としてステップS57以降（ステップS54乃至S56については後述する）の処理に進み、逆に、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 A_{thsd} を超えたか、或いは上記必須条件を満たしていないと判定した場合、当該フレームがCM候補にはならないとしてステップS50の処理に戻る。

【 0 1 0 5 】

上記ステップS52、S53の各判定処理の結果、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 A_{thsd} を超えず、且つ、上記必須条件を満たしていると判

定された場合、CM候補検出器19は、ステップS57の処理としてCM開始フレーム n_s を検索し、次に、ステップS58の処理としてCM終了フレーム n_e の検索を行い、更に、ステップS59の処理としてCM開始時刻 T_s を計算し、ステップS60としてCM長さ W を計算する。

【0106】

CM候補検出器19は、以上のステップS57乃至S60の検索及び計算を行った後、ステップS61において後述するCM候補テーブルを参照し、もし、CM開始時刻 T_s 及びCM長さ T_w の一致する候補がすでに当該CM候補テーブル中に存在するならば、そのまま再びステップS54乃至S56の処理に戻り、逆に存在しない場合には、新たなCM候補としてCM候補テーブルに追加した後、再びステップS54乃至S56の処理に戻る。

【0107】

ステップS54乃至S56では、全ての時間長に対して上述同様の処理を行った後、ステップS50に戻り、次の入力に対して同じ処理を繰り返すことを表している。

【0108】

なお、上記CM開始フレーム n_s とは、 n_1 , n_2 , n_3 で表される各フレームのうち時間条件に合致したフレームから、最新フレームの方向へ向かって、平均自乗振幅 $A[n]$ が自乗振幅の閾値 A_{thsd} を越える最初のフレーム番号である。また、CM終了フレーム n_e とは、0番目のフレームより過去の方方向に向かって、平均自乗振幅 $A[n]$ が自乗振幅の閾値 A_{thsd} を越えない最後のフレーム番号である。さらにCM開始時刻 T_s は、CM開始フレーム番号 n_s を用いて $T_s = T[n_s]$ として求められる。同様にCM長さ T_w は、 $T_w = T[n_e] - T[n_s]$ として求められる。

【0109】

ここで、図6に、上記必須条件の算出例を示す。この図6に示す $A[n]$ の項において、「o」は自乗振幅の閾値 A_{thsd} 未満の平均自乗振幅を持つフレームを示し、「x」は自乗振幅の閾値 A_{thsd} 以上の平均自乗振幅を持つフレームを示している。この例では、 $A[0]$, $C[0]$ 及び $A[n_1]$, $C[n_1]$ が条件を満たし、 n_1 より

左方で最初に $A[n] = x$ となるフレームが n_s 、0より右方に連続する最後の $A[n] = o$ となるフレームが n_e となる。

【0110】

以上の処理により、CM候補検出器19では、1フレーム（1サンプル）の特徴量が入力される毎にCM候補の検出を行い、CM候補が検出された場合にはCM候補テーブルにエントリーする。

【0111】

図7には、CM候補テーブルの構成例を示す。この図7において、CM候補テーブルの項目は、開始時刻 T_s 、長さ T_w 、及び後述する付加条件算出器20で算出する特徴量 Q_1 乃至 Q_{11} 、及び後述する付加条件判定器21で算出するスコア R とスコア判定結果 Z からなる。CM候補検出器19によるCM候補テーブル19aの段階では、開始時刻 T_s 、長さ T_w のみが記述される。このように、CM候補テーブルは、CM候補検出器19で得られるCM開始時刻 T_s 、長さ T_w と、付加条件算出器20で算出される特徴量 Q_1 乃至 Q_{11} と、付加条件判定器21で算出されるスコア R 及びスコア判定結果 Z とを記述し、それら特徴量を管理するための表である。また、CM候補テーブルは、そのエントリーがCMであるかないかの判定を受けるまで保持され、CMであると判断された場合には、後述するルール判定器22からCM検出出力4aとして出力され、CMでないと判断された場合には破棄される。

【0112】

上記CM候補検出器19により開始時刻 T_s 、長さ T_w のみが記述されたCM候補テーブル19aは、付加条件算出器20に送られる。

【0113】

付加条件算出器20では、CM候補テーブル19aにエントリーされた候補区間より、特徴量バッファ18を参照しながら、以下に示すような特徴量 Q_1 乃至 Q_{11} を抽出し、それをCM候補テーブル19aに追加記述し、CM候補テーブル20aとして付加条件判定器21に出力する。

【0114】

図8には、当該付加条件算出器20における特徴量 Q_1 乃至 Q_{11} の算出例を示

す。

【0115】

この図8において、横軸はフレーム番号（離散時間に相当）を表し、図8（A）はカットチェンジ検出出力 $C[n]$ 、図8（B）は放送モード信号 $2c$ の離散化数値 $B[n]$ 、図8（C）は音声信号の短時間離散スペクトル $S[k,n]$ 、図8（D）は音声信号の平均自乗振幅 $A[n]$ を表わし、 n_1 の間隔（図中点線で挟まれた区間）がCM候補である。なお、図8（A）において、図中CTで示す位置はカットチェンジ検出出力 $C[n]$ が1となっている位置（すなわちカットチェンジが検出された位置）を示している。また、図8（B）において、図中Mで示す区間はその区間が何らかの放送モードとなっていることを示している。図8の（C）において、図中 S_1 、 S_2 、 S_3 、 S_4 は何らかのスペクトル成分が存在することを示し、図8（D）において、図中AMは自乗振幅の変化を表している。また、図中 Q_1 から Q_{11} は、上記付加条件算出器20にて特徴量 Q_1 乃至 Q_{11} が計算される場所を示している。

【0116】

以下、付加条件算出器20で算出される各特徴量 Q_1 乃至 Q_{11} について個々に説明する。

【0117】

特徴量 Q_1 は前ブレーク長である。当該前ブレーク長とは、CM候補区間直前の小音量区間（前ブレーク区間と称する）、すなわち連続して $A[n]$ が所定の閾値 A_{thsd} 以下である時間長であり、図8中の一点鎖線で挟まれた区間長 BB が前ブレーク長 Q_1 である。

【0118】

特徴量 Q_2 は後ブレーク長である。当該後ブレーク長とは、CM候補区間直後の小音量区間（後ブレーク区間と称する）、すなわち連続して $A[n]$ が所定の閾値 A_{thsd} 以下である時間長であり、図8中の一点鎖線で挟まれた区間長 AB が後ブレーク長 Q_2 である。

【0119】

特徴量 Q_3 は前ブレーク最小振幅である。当該前ブレーク最小振幅 Q_3 は、前記

の前ブレイク区間における $A[n]$ の最小値である。

【0120】

特徴量 Q_4 は後ブレイク最小振幅である。当該後ブレイク最小振幅 Q_4 は、前記の後ブレイク区間における $A[n]$ の最小値である。

【0121】

特徴量 Q_5 は左右相関値である。当該左右相関値 Q_5 は、CM候補区間の音声の左右2チャンネルの音声信号 $S_L[m]$ 、 $S_R[m]$ の相関値である。これは、式(5)乃至式(7)の $A_{LR}[n]$ 、 $A_{LL}[n]$ 、 $A_{RR}[n]$ を利用して、式(14)に従って算出することができる。

【0122】

【数7】

$$Q_5 = \frac{\sum_{n=n_s}^{n_e-1} A_{LR}[n]}{\sum_{n=n_s}^{n_e-1} A_{LL}[n] \sum_{n=n_s}^{n_e-1} A_{RR}[n]} \quad (14)$$

【0123】

この式(14)の演算では、フレームのオーバーラップにより原波形が部分的に複数回加算されることになるが、そのことはこのシステムに実質的な影響は及ばさない。また、原波形をそのまま保持できるだけのメモリ容量及び処理速度がある場合には、この演算は原波形の相互相関と置き換えることもできる。

【0124】

特徴量 Q_6 は平均振幅値である。当該平均振幅値 Q_6 は、CM候補区間の音声信号の振幅のRMS値(平均自乗振幅)である。これは、式(15)により計算することができる。

【0125】

【数8】

$$Q_6 = \sqrt{\frac{1}{n_e - n_s} \sum_{n=n_s}^{n_e-1} A[n]^2} \quad (15)$$

【0126】

この式(15)の演算では、上記左右相関演算の場合と同様に、フレームのオーバーラップ次第では原波形が部分的に複数回加算されることになるが、そのことは実質的な影響を及ぼさない。また、原波形をそのまま保持できるだけのメモリ容量及び処理速度がある場合には、この演算は原波形のRMS演算と置き換えることもできる。

【0127】

特徴量 Q_7 はカット数である。当該カット数 Q_7 は、CM候補区間中に存在するカットチェンジの回数(上記CTの数)を数える演算となる。すなわちこれは、 $[n_s, n_e)$ の区間で $C[n] = 1$ となる回数を数える演算となる。

【0128】

特徴量 Q_8 は放送モードである。ここでの放送モードは、CM候補区間中で最も支配的な放送モードのことである。これは、 $[n_s, n_e)$ の区間の $B[n]$ 値の中で、最も頻発する放送モード Q_8 を選ぶ演算である。

【0129】

特徴量 Q_9 は隣接候補数である。当該隣接候補数 Q_9 は、あるCM候補に対して、その前後にある有音区間もCM候補であるかどうかを表わし、両側ともCM候補であれば「2」、片側のみCM候補であれば「1」、どちらもCM候補でなければ「0」の値をとる。この演算は、CM候補テーブルを検索することで行われ、開始時刻 T_s と長さ T_w と後ブレーク長 Q_2 の和($T_s + T_w + Q_2$)が、他のCM候補の開始時刻(T'_s)と一致するかどうかで後側候補の判定が行われる。同様に、開始時刻 T_s と前ブレーク長 Q_1 の差($T_s - Q_1$)が、他のCM候補の開始時刻 T'_s と長さ T'_w の和($T'_s + T'_w$)と一致するかどうかで、前側候補の判定が行われる。

【0130】

特徴量 Q_{10} 、 Q_{11} はスペクトル差分エネルギーである。当該スペクトル差分エネルギー Q_{10} 、 Q_{11} は、CMと番組本編やCMと他のCMとの境界での音質変化を定量化するために用いられる。これは、上記境界の両側における平均スペクトルの差の自乗和として定義され、式(16)乃至(21)に従って計算される。

【0131】

【数 9】

$$S_1[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_e - n] \quad (16)$$

$$S_2[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_s + n] \quad (17)$$

$$S_3[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_e - n] \quad (18)$$

$$S_4[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_s + n] \quad (19)$$

$$Q_{10} = \frac{1}{S_{norm}^2} \sum_k \left(S_2[k] - S_1[k] \right)^2 \quad (20)$$

$$Q_{11} = \frac{1}{S_{norm}^2} \sum_k \left(S_4[k] - S_3[k] \right)^2 \quad (21)$$

【0132】

但し、式中のNはスペクトルの平均をとるフレーム数、 n'_e はCM候補区間の直前の有音区間の終了フレーム番号（図8参照）、 n'_s はCM候補区間の直後の有音区間の開始フレーム番号、 $S_1[k]$ はCM候補区間の直前の有音区間の終了直前の平均スペクトラム、 $S_2[k]$ はCM候補区間開始直後の平均スペクトラム、 $S_3[k]$ はCM候補区間終了直前の平均スペクトラム、 $S_4[k]$ はCM候補区間の直後の有音区間の開始直後の平均スペクトラム、 S_{norm} は適切な規格化定数である。

【0133】

上記付加条件算出器20は、以上により算出した Q_1 から Q_{11} までの特徴量を、CM候補テーブル19aに追加記述し、CM候補テーブル20aとして出力する。当該CM候補テーブル20aは、付加条件判定器21に送られる。

【0134】

付加条件判定器21は、CM候補テーブル20aを入力とし、CM候補の特徴量を、閾値関数などにより非線型にパラメータ変換した後、荷重加算することでCM候補に対するスコアRを算出し、Rが所定の閾値以上である場合には有力CM候補と判定する。付加条件判定器21は、これらスコアRとスコア判定結果ZをCM候補テーブル20aに追加記述し、CM候補テーブル21aとして出力する。

【0135】

図9には、付加条件判定器21の概略構成を示す。

【0136】

この図9において、CM候補テーブル21aの各特徴量 Q_1 乃至 Q_L は、それぞれ対応する関数演算器50₁乃至50_Lに送られ、それぞれ対応するパラメータ変換関数 $H_1()$ 乃至 $H_L()$ による変換演算が施された後、さらにそれぞれ対応する重み付け器51₁乃至51_Lにより荷重 W_1 乃至 W_L との積がとられる。各重み付け器51₁乃至51_Lにより重み付けがなされた後の特徴量は、総和加算器52での総和加算によりスコアRが算出される。この総和加算器52から出力されたスコアRは、スコア判定器53にて所定の閾値と比較され、スコアRが所定の閾値以上である場合には有力CM候補である旨を示す判定結果が出力される。なお、スコア判定器53によるスコア判定により所定の閾値未満であると判定されたCM候補は、テーブルから消去される。

【0137】

より具体的に説明すると、当該付加条件判定器21におけるスコア算出演算は、式(22)に従って行われる。

【0138】

【数10】

$$R = \sum_{i=1}^L W_i H_i(Q_i) \quad (22)$$

【0139】

ただし、 $H_1()$ は各特徴量に対して予め定めるパラメータ変換関数、 W_1 は予め決定しておく荷重、 L は特徴量数(=11)である。なお、 1 は1乃至11のうちの任意の数である。

【0140】

ここで、各関数演算器50₁乃至50_Lにおけるパラメータ変換関数 $H_1()$ は、最も簡単には矩形関数でよい。すなわち例えば、図10(B)に示すような矩形関数 $\text{Rect}(x; t_1, t_2)$ を用い、予め各特徴量について決定しておく標準値の下上限値を t_{11} 、 t_{21} とし、式(23)により例えば Q_1 が標準値の範囲内ならば1

、範囲外ならば0となるようにする。

【0141】

$$H_1(Q_1) = \text{Rect}(Q_1; t_1, t_2) \quad (23)$$

なお、前記境界付近で滑らかに0から1、1から0へ推移させるようにする場合には、例えば式(24)のような、シグモイド関数 $\text{Sigm}(x; t_1, t_2)$ を用いることもできる。

【0142】

【数11】

$$H_1(Q_1) = \text{Sigm}(Q_1; t_{11}, t_{21}) = \frac{1}{1 + \exp(-(x - t_{11})/\sigma_{11})} \cdot \frac{1}{1 + \exp((x - t_{21})/\sigma_{21})} \quad (24)$$

【0143】

図10(C)にシグモイド関数の概形を示す。ただし、 s_{11}, s_{21} は推移の程度を表わす定数であり、予め特徴量の分布などに基づき決定しておく。

【0144】

また、上記各重み付け器51₁乃至51_Lによる加算荷重 W_1 は、予め特徴量の統計的性質に基づき人為的に決定しておくこともできるが、既知の学習サンプルに対して、ニューラルネットワーク（例えば中川著「パターン情報処理」丸善（1999）などに詳説）の要領で学習することで、自動的に荷重を決定することも可能である。なお、1は1乃至11のうちの任意の数である。

【0145】

さらに、上記判定器53におけるスコア判定は、式(25)のように、スコアRの閾値処理により行う。

【0146】

$$Z = \text{Unit}(R - t_r) \quad (25)$$

ただし、 $\text{Unit}(x)$ は、図10(A)に示すように、 $x > 0$ で1、 $x < 0$ で0となる単位ステップ関数であり、 t_r は予め定めるか或いは学習により自動的に決まる判定閾値である。

【 0 1 4 7 】

次に、ルール判定器 2 2 は、上記付加条件判定器 2 1 でのスコア判定により得られた CM 候補テーブル 2 1 a を入力とし、後述するような所定のルール判定により最終的な CM 検出出力 4 a として CM 開始時刻と長さを出力する。すなわち、当該ルール判定器 2 2 では、同一時刻に複数の CM 候補があった場合（以下、競合関係という）に、どちらがより CM として確からしいかをルール処理により判定する。

【 0 1 4 8 】

以下、ルール判定器 2 2 の動作を図 1 1 のフローチャートを用いて説明する。

【 0 1 4 9 】

まず、ルール判定器 2 2 は、ステップ S 7 0 として、CM 候補テーブルより、判定すべき CM 候補を選択する。この選択は、CM 候補テーブル中で最古の候補であり、予め設定された時間 T_4 が経過したものから順に行われる。 T_4 は、数個の CM が十分含まれる程度の時間長であり、例えば 1 5 0 秒間とする。

【 0 1 5 0 】

続いて、ルール判定器 2 2 は、ステップ S 7 1 として、選択した CM 候補の区間中（ T_s から $T_s + T_w$ までの間）に、他の CM 候補が存在するかどうか、CM 候補テーブル中を検索する。このステップ S 7 1 において、他の CM 候補が存在しないと判定した場合（No）、この CM 候補は CM 検出出力として出力され、CM 候補テーブルより消去される。

【 0 1 5 1 】

一方、ステップ S 7 1 において、他の CM 候補が存在すると判定された場合（Yes）、それらは競合関係にあるとして、ステップ S 7 2 にてまず最小長さ優先ルールが適用される。ここで、最小長さ優先ルールは、ある時区間が、複数の長さの異なる CM 候補の組み合わせにより構成され得る場合、より短い CM 候補で構成されている方を優先するというルールである。すなわち、例えば 3 0 秒という時区間に対して、1 つの 3 0 秒 CM という可能性と、2 つの 1 5 秒 CM の組み合わせという可能性の、両方が同時に候補として存在する場合には、1 5 秒 CM を選択し、3 0 秒 CM を棄却するというルールである。

【 0 1 5 2 】

図 1 2 を用いて、この最小長さ優先ルールの一例を説明する。

【 0 1 5 3 】

なおこの例には、図 1 2 (A) のように、実際には 4 つの CM 1 乃至 CM 4 が連続して放送されている区間に対し、図 1 2 (B) 中 A 乃至 H で示すような 8 つの候補が CM 候補テーブルに存在する場合が示されている。

【 0 1 5 4 】

先ず、図 1 2 (C) に示すように、A の CM 候補が判定中であるとすると、この候補 A と競合する候補は E と H である。しかしながら、E の区間は A と B で記述でき、また、H の区間は A と B と C と D 等で記述できることから、それぞれ棄却され、A が採用される。続いて、図 1 2 (D) に示すように、B が判定中となったときには、F が競合相手となる（このとき、E、H は A の判定により棄却済みとなっている）が、F の区間は B と C で記述できることから棄却され、B が採用される。同様に、図 1 2 (E) に示すように、C が判定中の場合には、G が競合相手となるが、G の区間は C と D で記述されることから棄却され、C が採用される。最後に、図 1 2 (F) に示すように、D が判定されるときには、すでに競合相手は存在しないので、そもそもこのルールを適用する必要はなく、当該 D がそのまま採用される。

【 0 1 5 5 】

以上により、この時区間からは、CM 候補として A、B、C、D が選択されることとなる。このルールが適用できない競合関係については、そのまま CM 候補テーブルに残してこの処理を終了する。

【 0 1 5 6 】

図 1 1 に戻り、ステップ S 7 2 のルール判定の後、ルール判定器 2 2 の処理は、ステップ S 7 3 に進む。ステップ S 7 3 に進むと、ルール判定器 2 2 は、最小長さ優先ルールを適用した結果、判定中の CM が棄却されたか否か判断する。このステップ S 7 3 において、判定中の CM が棄却されと判断した場合（Y e s）、ルール判定器 2 2 は、その候補を CM 候補テーブルから消去し、ステップ S 7 0 に戻る。一方、ステップ S 7 3 において、判定中の CM が棄却されていないと

判断した場合（N o）、ルール判定器 2 2 は、ステップ S 7 4 において、再び判定中の CM 候補の区間中に他の CM 候補が存在するかどうか、テーブル中を検索する。

【 0 1 5 7 】

このステップ S 7 4 において他の CM 候補が存在しないと判定した場合（N o）、ルール判定器 2 2 は、ステップ S 8 0 において判定中の CM 候補を CM 検出出力から出力し、CM 候補テーブルから消去する。一方、ステップ S 7 4 にて他の CM 候補が存在すると判断した場合（Y e s）、ルール判定器 2 2 の処理は、ステップ S 7 5 に進む。

【 0 1 5 8 】

ステップ S 7 5 に進むと、ルール判定器 2 2 は、隣接優先ルールを適用する。ここで、隣接優先ルールとは、複数の CM 候補が競合関係にある場合、それぞれ直前又は直後に隣接する CM 候補を検索し、それが存在する方を優先するというルールである。

【 0 1 5 9 】

図 1 3 を用いて、当該隣接優先ルールについて説明する。

【 0 1 6 0 】

なおこの例には、図 1 3（A）のように、実際には 4 つの CM 1 1 乃至 CM 1 4 が連続して放送されている区間に対し、図 1 3（B）中 I 乃至 N で示するような 6 つの候補が存在する場合が示されている。また、この例の場合、候補 M 及び N は、偶然 CM 中にカットチェンジや小音量区間が存在したために候補となっているが、このような候補は、実際には誤った区間であるとはいえ、内容的には CM を含んでいるため、CM らしさを判定する付加条件のスコア判定によっても、棄却されない場合があるものである。

【 0 1 6 1 】

このような例において、まず、図 1 3（C）に示すように、最古の I が判定される候補となる。当該 I と競合するものとして M があるが、I には隣接する候補 J が存在するのに対し、M には隣接する候補がないため、I を採用し、M を棄却する。次に、図 1 3（D）に示すように、J が判定される候補となった場合、J

と競合する候補としてNがあるが、Jには隣接する候補I、Kが存在するのに対し、Nには存在しないため、Jが採用されNが棄却される。次に、図13(E)、(F)に示すように、残りの候補K、Lには、既に競合する候補がなくなるため、このルールは適用されず、これらK、Lがそのまま採用される。

【0162】

以上により、この図13に例示した区間からは、I、J、K、LがCM候補として選択されることとなる。

【0163】

なお、競合関係の候補のいずれにも隣接候補が無い場合、及び複数の候補にそれぞれ隣接候補がある場合には、それらはどちらも棄却されず、CM候補テーブルに残される。

【0164】

図11に戻り、ステップS75の処理後、ルール判定器22の処理は、ステップS76に進む。ステップS76に進むと、ルール判定器22は、隣接優先ルールを適用の結果、判定中のCMが棄却されたか否か判断する。このステップS76において、判定中のCMが棄却されたと判断した場合(Yes)、ルール判定器22は、その候補をCM候補テーブルから消去し、ステップS70の処理に戻る。一方、ステップS76において棄却されていないと判定された場合(No)、ルール判定器22は、次のステップS77において、再び判定中のCM候補の区間中に、他のCM候補が存在するかどうか、CM候補テーブル中を検索する。

【0165】

このステップS77において、他のCM候補が存在しないと判定された場合(No)、ルール判定器22は、ステップS80において、判定中のCM候補をCM検出出力から出力し、CM候補テーブルから消去する。一方、ステップS77において、他のCM候補が存在すると判定した場合(Yes)、ルール判定器22は、ステップS78において、スコア優先ルールを適用する。ここで、スコア優先ルールとは、上記の各ルールによっても競合関係が解消されない場合、付加条件判定器21により得られた判定スコアRの高い候補を優先するというルールである。このスコア優先ルールは、対象となる競合関係が解消するまで繰り返し適

用する。

【0166】

図14を用いて、当該スコア優先ルールについて説明する。

【0167】

なおこの例には、図14(A)のように、実際には4つのCM21乃至CM24が連続して放送されている区間に対し、図14(B)中P乃至Wで示すような7つの候補が存在する場合が示されている。

【0168】

この例において、先ず、図14(C)に示すように、最古のPが判定される候補となるが、この候補PはUと競合関係がある。但し、このときの競合関係は、前記最小長さ優先ルールによっても、また、隣接優先ルールによっても競合が解消されない。

【0169】

したがって、この場合には、これら競合関係にある候補と関連する全ての競合関係を、CM候補テーブル中から検索する。すなわち、この場合は、(P-U)、(U-Q)、(Q-V)、(V-R)、(R-W)、(W-S)という、7候補に対する6つの競合関係が全て関連しているので、スコア優先ルールでは、これら関連する候補の中で最もスコアの高い候補を採用する。この例の場合、判定スコアR(2.0)が最も高いスコアであるため、図14(D)に示すように、このスコアが採用され、その結果、Rと競合関係にある候補V、Wは棄却される。

【0170】

しかしながら、図14(E)に示すように、これによっても(P-U)の競合関係は解消されていない。したがって、再びこれらと関連する全ての競合関係を、CM候補テーブル中から検索する。今回は、Vが棄却されたことにより、(P-U)、(U-Q)という、3つの候補が関係する2つの競合関係のみとなる。

【0171】

さらに、これらの候補の中で最もスコアの高い候補Q(1.9)を採用し、図14(F)に示すように、Qと競合関係にある候補Uを棄却する。

【0172】

以上によって、Pに関係する競合関係はなくなり、Pが採用される。また、U、V、Wは全て棄却され、Q、R、Sが採用されることとなる。

【0173】

なお、仮に、関連する全ての競合関係を検索せず、対象となる競合関係（この例の場合、P、U）のみでスコア優先ルールを適用すると、先ずUが採用され、Pは棄却される。後にUとQとの競合関係により、一時採用されたUもまた棄却されてしまう。このように、ルール判定器22では、偶然の処理順序により候補Pが棄却されるようなことのないよう、関連競合関係の検索を行っている。

【0174】

以上のスコア優先ルールにより、選択された候補に関する競合関係は必ず解消されることになる。

【0175】

図11に戻り、ステップS78の処理後、ルール判定器22の処理は、ステップS79に進む。ステップS79に進むと、ルール判定器22は、スコア優先ルールを適用の結果、判定中の候補が棄却されたか否か判断する。このステップS79において、判定中の候補が棄却されたと判断した場合（Yes）、ルール判定器22は、その候補をCM候補テーブルより消去し、ステップS70に戻る。一方、ステップS79において、棄却されなかった場合、ルール判定器22は、ステップS80のCM検出出力として、開始時刻とその長さを出力し、CM候補テーブルから消去した後、ステップS70に戻る。

【0176】

以上説明したように、本実施の形態の第1の具体例のCM検出部4においては、ほぼ全てのCMが満足する必須条件に基づき、決定論的に番組中からCMの候補を抽出し、CMらしさの指標である付加条件に基づく特徴量の統計論的な評価により候補を選択し、論理条件により候補のオーバーラップ関係を解消することで、精度の高いCM検出を実現している。また、本実施の形態では、例えば現行のアナログTV放送の映像音声記録装置を例にとったが、デジタルTV放送等に適用される場合にも同様のCM検出部4が適用できることは明らかである。また、例えばラジオ放送に適用される場合には、上記CM検出部4から映像信号の

処理を担当する部分を省略することで同様の機能が実現できる。

【0177】

次に、本発明の第2の具体例としてのCM検出部4について以下に説明する。

【0178】

図15には、本発明の第2の具体例としてのCM検出部4の詳細な構成を示す。

【0179】

当該第2の具体例のCM検出部4は、前述した付加条件の中でも基本的なもののみを実装するようにしており、前述した1)乃至14)の付加条件のうち、11)乃至14)に関しては導入しないことで、装置構成を簡略化している（複雑になることを防いでいる）。

【0180】

この第2の具体例のCM検出部4も、図2の例と同様にフロントエンド部とバックエンド部とから構成されている。なお、この図15において、図2の各構成要素と同じ動作を行う部分については、同一の指示符号を付して、それらの説明は省略する。

【0181】

以下、図15の構成において、新たに追加された各構成要素（101、102、103）と、付加条件算出器20において新たに拡張された機能についてのみ説明する。

【0182】

フロントエンド部に設けられた音源識別器101は、デジタル化及びフレーム化された音声信号2bを入力とし、この音声信号2bの該当フレームに関する音源名を出力する。音源名としては、例えば、音声、音楽、音声と音楽、その他を挙げることができる。なお、入力された音声信号の音源識別を実現する技術としては、例えば、河地、他による、「VQ歪みに基づく放送音の自動分類」信学技報、DSP97-95/SP97-50、43/48(1998)に記載された技術や、南、他による、「音情報を用いた映像インデクシングとその応用」信学論、Vol.J81-D-II、No.3、529/537(1998)に記載された技術、安部による、特願平11-190693号の

明細書及び図面に記載された技術などがあり、これらを利用することができる。

【0183】

この音源識別器101により識別された各音源名は、例えば音声＝1、音楽＝2、などのように、各フレーム毎に適切に数値化され、特徴量 $U[n]$ として特徴量バッファ18に入力される。

【0184】

フロントエンド部に設けられた番組ジャンルデータ又は番組ジャンル識別器102は、現在処理している番組のジャンル名を出力するものである。番組ジャンルは、例えば、ニュース、ドラマ、野球、サッカーなどである。番組ジャンルデータは、テレビ番組表などから入力してもよく、また近年ではインターネット等を通じて自動的に取得することもできる。または、外部情報に頼らず音声及び映像信号から番組ジャンルを識別する装置を用いることも可能である。なお、音声及び映像信号から番組ジャンルを識別する技術としては、例えば安部による、特願平11-190693号の明細書及び図面に記載された技術などを利用することができる。

【0185】

この番組ジャンルデータ又は番組ジャンル識別器102により分類された番組ジャンル名は、例えばニュース＝1、ドラマ＝2、などのように、各フレーム毎に適切に数値化され、特徴量 $W[n]$ として特徴量バッファ18に入力される。

【0186】

フロントエンド部のその他の各構成要素は、図2の例と同様である。

【0187】

この第2の具体例の場合、フロントエンド部に、上記音源識別器101と番組ジャンルデータ又は番組ジャンル識別器102を設け、これらにより得られた各特徴量 $U[n]$ と $W[n]$ を特徴量バッファ18に蓄積することで、当該特徴量バッファ18においては、式(10)に示した特徴量 $G[n]$ が、式(26)のように拡張されることになる。

【0188】

$$G[n] = \{C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n],$$

$$B[n], T[n], U[n], W[n] \quad (26)$$

バックエンド部のCM検出器19は、前述の図2のものと同様のものであるが、当該第2の具体例の場合、CM候補テーブル19a乃至21aは、次のように拡張される。すなわち、この第2の具体例の場合のCM候補テーブル19a乃至21aは、前述した Q_1 から Q_{11} までの特徴量に加え、図16に示すように、後述する特徴量 Q_{12} から Q_{15} が拡張される。なお、図16は、 Q_1 から Q_{11} までの特徴量についての図示を省略している。

【0189】

また、バックエンド部のCM確率データベース103には、予め、時間帯に応じたCMの放送確率、及び、番組ジャンルと経過時間に応じたCMの放送確率をデータとして蓄積してある。このCM確率データベース103からは、現在時刻に応じてそれらの確率が読み出され、付加条件算出器20に入力するようになされている。なお、これらの確率のデータは、実際の放送を元に統計をとることで作成することができる。

【0190】

この第2の具体例の場合の付加条件算出器20は、前述の特徴量 Q_1 から Q_{11} に加え、次の特徴量 Q_{12} から Q_{15} の演算を行うよう拡張される。

【0191】

ここで、特徴量 Q_{12} は、CM候補区間中に、音声区間があったかどうかを検出して求められるものである。音声の有無を表す特徴量 Q_{12} は、式(27)に従って検出される。

【0192】

【数12】

$$Q_{12} = \begin{cases} 1 & \left(\text{if } \exists U[n] = (\text{音声}) \text{ or } \exists U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e \right) \\ 0 & (\text{otherwise}) \end{cases} \quad (27)$$

【0193】

特徴量 Q_{13} は、上記音声の有無と同様に、CM候補区間中に、音楽区間があったかどうかを検出して求められるものである。この音楽の有無を表す特徴量 Q_{13} は、式(28)に従って検出される。

【0194】

【数13】

$$Q_{13} = \begin{cases} 1 & (\text{if } \exists U[n] = (\text{音声}) \text{ or } \exists U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e) \\ 0 & (\text{otherwise}) \end{cases} \quad (28)$$

【0195】

特徴量 Q_{14} は、現在時刻に応じたCMの発生確率（時間帯確率）である。付加条件算出器20では、CM確率データベース103より提供されるCMの放送確率を、そのまま特徴量 Q_{14} に代入する。

【0196】

特徴量 Q_{15} は、番組ジャンル及びその番組の開始からの経過時間に従うCMの放送確率（番組ジャンル確率）である。付加条件算出器20では、CM確率データベース103より提供されるCMの放送確率を、そのまま特徴量 Q_{15} に代入する。

【0197】

付加条件判定器21以降は、変数としての特徴量 Q_{12} 乃至 Q_{15} が拡張されるだけであり、前述の図2のCM検出部4の場合と同様であるため、説明を省略する。

【0198】

この場合のCM検出部4においては、以上の拡張により、放送信号の音源に応じたCM検出を行うことができ、また、現在時間に応じたCM検出を行うこと、さらに、番組ジャンルに応じたCM検出を行うことが可能となる。

【0199】

CM検出部4の第3の具体例として、例えば、図17に示すように、小振幅回

数、小振幅区間、および信号分散を、それぞれ特徴量 Q_{16} 乃至 Q_{18} として、付加条件算出器20により算出させるようにすることができる。

【0200】

小振幅回数とは、音声信号の振幅が、予め設定されている所定の閾値を下回る回数を意味する。付加条件算出器20は、例えば、図18のフローチャートに示すような処理を行うことで、小振幅回数を計算する。

【0201】

最初にステップS90において、カウンタCとフラグFがリセットされる。カウンタCには、小振幅回数が保持され、フラグFは、小振幅区間であることを表す。ステップS90においては、さらに、時刻nがCM候補の開始時刻に設定される。

【0202】

ステップS91において、音声信号の振幅信号A[n]が取得される。nは、離散化された時刻に対応する。

【0203】

ステップS92において、いまフラグFがセットされているか否かが判定され、セットされていない場合、ステップS93に進み、ステップS91で取得された信号振幅A[n]が、予め設定されている所定の閾値A1より小さいか否かが判定される。音声信号の振幅A[n]の値が閾値A1より等しいか、それより大きいと判定された場合、ステップS97に進み、時刻nの値がインクリメントされる。そして、ステップS98において、時刻nの値がCM候補の終了時刻に達したか否かが判定され、終了時刻に達していない場合には、ステップS91に戻り、次の時刻のタイミングにおける振幅A[n]が取得される。

【0204】

以上のような処理がステップS93において、振幅A[n]の値が閾値A1より小さいと判定されるまで繰り返し実行される。ステップS93において、振幅A[n]の値が閾値A1より小さいと判定された場合、ステップS94に進み、カウンタCの値が1だけインクリメントされ、かつ、フラグFがセットされる。

【0205】

その後、ステップS97に進み、時刻 n がインクリメントされ、ステップS98において、インクリメントされた時刻 n の値がCM候補の終了時刻に達しているか否かが判定され、まだ達していない場合には、ステップS91に戻り、次のタイミングの振幅 $A[n]$ が取得される。

【0206】

そして、ステップS92において、フラグFがセットされているか否かが判定され、いまの場合、フラグFがセットされているので、ステップS95に進み、ステップS91で取得された振幅 $A[n]$ の値が、予め設定されている閾値 $A2$ より大きいかが判定される。なお、この閾値 $A2$ の値は、ステップS93において比較される閾値 $A1$ より大きい値($A2 > A1$)とされている。

【0207】

ステップS95において、振幅 $A[n]$ の値が閾値 $A2$ より大きくないと判定された場合、ステップS97に進み、時刻 n の値がインクリメントされる。

【0208】

ステップS98において、時刻 n の値がCM候補の終了時刻に達しているか否かが再び判定され、達していない場合には、ステップS91に戻り、次のタイミングの振幅 $A[n]$ が取得される。

【0209】

ステップS92において、フラグFがセットされているか否かが再び判定され、いまの場合、まだセットされているので、ステップS95に進み、取得された振幅 $A[n]$ が閾値 $A2$ より大きくないと判定された場合、ステップS97に進み、上述した場合と同様の処理が繰り返し実行される。

【0210】

以上のようにして、振幅 $A[n]$ の値が、より小さい閾値 $A1$ より小さいと判定された場合、カウンタCの値が1だけインクリメントされた後、振幅 $A[n]$ の値が、閾値 $A1$ より若干大きい値の閾値 $A2$ より大きくなるまで待期する。

【0211】

ステップS95において、振幅 $A[n]$ の値が閾値 $A2$ より大きいと判定された場合、ステップS96に進み、フラグFがリセットされる。その後ステップS

97に進み、時刻 n の値がインCREMENTされる。ステップ S98において、時刻 n の値がCM候補の終了時刻に達したか否かが判定され、達していない場合には、ステップ S91に戻り、それ以降の処理が繰り返し実行される。

【0212】

ステップ S98において、時刻 n の値がCM候補の終了時刻に達したと判定された場合、処理は終了される。

【0213】

ステップ S95において基準とされる閾値 A2の値を、ステップ S93の処理において設定される閾値 A1より大きく設定することで、判定処理に、いわゆるヒステリシス特性を持たせることが可能となる。すなわち、振幅 $A[n]$ の値が、より小さい閾値 A1より小さくなったとき、小振幅区間に入ったと判定されるが、振幅 $A[n]$ の値が、閾値 A1より若干大きくなっても、閾値 A2より小さい場合には、まだ小振幅期間中であると判定され、閾値 A1より大きい閾値 A2よりさらに大きくなったとき、初めて、小振幅期間が終了したと判定される。これにより、小振幅期間中のわずかな振幅の変化に起因して、小振幅回数が必要以上に大きな値にカウントされることが防止される。

【0214】

以上のようにして、CM候補の期間における小振幅の回数がカウンタ C に設定され、このカウンタ C の値が特徴量 Q_{16} として出力される。

【0215】

図19は、小振幅回数の具体的な計測結果の例を表している。図19における横軸は、小振幅回数を表し、縦軸は、相対度数を表している。図19(A)のグラフは、実験データから得られた490個のCM候補のうち、実際にCMであった352個の度数分布を表しており、図19(B)は、そのうちのCMでなかった138個の度数分布を表している。すなわち、図19(B)は、本編中で偶然、音量条件や映像条件が満たされたためにCM候補として検出されたものである。

【0216】

これらの図を比較して明らかなように、CMである場合(図19(A))、小

振幅回数は、0回から2回に集中するのに対して、CMでない場合には（図19（B））、小振幅回数は、7回乃至9回と多くなることがわかる。

【0217】

次に、図20のフローチャートを参照して、CM検出部4の付加条件算出器20が実行する小振幅区間長計算処理について説明する。最初に、ステップS110において、小振幅区間長を表すカウンタDの値がリセットされ、かつ、時刻nの値がCM候補の開始時刻にセットされる。

【0218】

次にステップS111において、音声信号の振幅A[n]が取得され、ステップS112において、ステップS111で取得された振幅A[n]の値が、予め設定されている所定の閾値A1より小さいか否かが判定される。この閾値A1は、図18のステップS93における閾値A1と等しい値とされているが、異なる値とすることも可能である。

【0219】

ステップS112において、振幅A[n]の値が閾値A1と等しいか、それより大きいと判定された場合、ステップS114に進み、時刻nの値がインクリメントされる。そして、ステップS115において、インクメントされた時刻nの値が、CM候補の終了時刻に達したか否かが判定され、終了時刻に達していない場合には、ステップS111に戻り、次のタイミングの振幅A[n]が取得される。

【0220】

そして、その振幅A[n]の値が、ステップS112において、閾値A1より小さいか否かが再び判定され、振幅A[n]の値が閾値A1より小さくない場合には、ステップS114に進み、時刻nの値がさらにインクメントされる。

【0221】

以下、同様の処理が繰り返し実行され、ステップS112において、振幅A[n]の値が、閾値A1と等しいか、それより大きいと判定された場合、ステップS113に進み、カウンタDの値が1度だけインクメントされる。その後、ステップS114に進み、時刻nの値がインクリメントされる。ステップS115に

において、時刻 n の値が、CM候補の終了時刻に達したか否かが判定され、達していない場合には、ステップ S 1 1 1 に戻り、次のタイミングの振幅 $A[n]$ が取得され、ステップ S 1 1 2 において、その振幅 $A[n]$ の値が、閾値 A_1 より小さいか否かが再び判定される。振幅 $A[n]$ の値が閾値 A_1 より小さい場合には、ステップ S 1 1 3 において、カウンタ D の値が再び 1 だけインクリメントされる。

【0222】

以上のような処理が繰り返し実行されることで、カウンタ D の値は、振幅 $A[n]$ の値が閾値 A_1 より小さい期間に対応する値となる。

【0223】

ステップ S 1 1 5 において、時刻 n の値が CM 候補の終了時刻に達したと判定された場合、ステップ S 1 1 6 に進み、カウンタ D の値が規格化される。すなわち、カウンタ D の値は、サンプリング周波数 f_s で割算されることで、規格化され、その値が特徴量 Q_{17} として出力される。

【0224】

図 2 1 は、図 1 9 における場合と同様に、490 個の CM 候補のうちの 352 個の実際の CM と、138 個の CM でなかった場合の小振幅区間長の総和を表している。図 2 1 において、横軸は、小振幅区間長の総和（単位は秒）を表し、縦軸は、相対度数を表している。図 2 1 (A) と図 2 1 (B) を比較して明らかに、CM の小振幅区間長の総和は、20 ms 程度以下に集中している（図 2 1 (A)）のに対して、CM でない場合には、1.0 s 以上の長さに集中している（図 2 1 (B)）。

【0225】

さらに、付加条件算出器 20 は、音声信号の振幅の分散を式 (29) に基づいて演算する。この式 (29) において、 s は、CM 候補の離散開始時刻を表し、 e は、CM 候補の離散終了時刻を表し、 v は、信号の分散を表す。付加条件算出器 20 は、この値 v をそのまま特徴量 Q_{18} として出力するか、あるいはその平方根を取って、標準偏差を特徴量 Q_{18} として出力する。あるいはまた、付加条件算出器 20 は、標準偏差を平均値で割算し、相対標準偏差を特徴量 Q_{18} として出力

することができる。

【0226】

【数14】

$$v = \frac{1}{e-s} \sum_{n=s}^{e-1} A^2[n] - \left(\frac{1}{e-s} \sum_{n=s}^{e-1} A[n] \right)^2 \quad (29)$$

【0227】

図22は、振幅の分散の例を表している。図22(A)は、490個のCM候補のうちの、352個の実際にCMであった場合の振幅の分散を表しており、図22(B)は、138個のCMでなかった場合の分散を表している。なお、図22において、横軸は、相対標準偏差を表しており、縦軸は、相対度数を表している。

【0228】

これらの図を比較して明らかなように、CMの場合(図22(A))、振幅の相対標準偏差がほぼ0.6以下に集中しているのに対して、CMでない場合には(図22(B))、0.7以上となることが多いことがわかる。

【0229】

従って、小振幅回数、小振幅区間長、および振幅分散を特徴量として利用することで、より正確にCMを検出することが可能となる。

【0230】

次に、本発明の第2の実施の形態としての映像音声記録装置について以下に説明する。

【0231】

図23には、第2の実施の形態の映像音声記録装置の概略構成を示す。

【0232】

なお、この図23において、図1の各構成要素と同じ動作を行う部分については、同一の指示符号を付して、それらの説明は省略する。また、第2の実施の形態の映像音声記録装置の場合のCM検出部4は、前記第1の具体例、第2の具体例、および第3の具体例の何れも適用できる。

【 0 2 3 3 】

以下、図 2 3 の構成において、新たに追加された各構成要素（1 1 0， 1 1 1）と、CM検出部 4 において新たに拡張された機能についてのみ説明する。

【 0 2 3 4 】

先ず、この第 2 の実施の形態の映像音声記録装置における CM 検出部 4 は、前述のように、式（1 0）に示される各 CM 候補の特徴量 $G[n]$ を、内部で算出している。また、当該第 2 の実施の形態の場合、CM 検出部 4 は、最終的に CM として検出されたものに関して、その開始時刻及び時間長と共に、CM 開始フレーム $n = n_s$ から終了フレーム $n = n_e$ に渡って、 $G[n]$ を CM データベース 1 1 0 に出力するように機能が拡張されている。

【 0 2 3 5 】

CM データベース 1 1 0 は、上記検出された CM に関して、その開始時刻、時間長、特徴量 $G[n]$ を保存する。

【 0 2 3 6 】

CM 特徴量比較器 1 1 1 は、ユーザが入力する検索指令に基づき、データベース 1 1 0 に保存されている全て又は一部の CM から、ユーザが指定した CM と同じ CM を抽出し、CM 検索出力 1 1 1 a として出力する。

【 0 2 3 7 】

この CM 特徴量比較器 1 1 1 の動作を、図 2 4 を用いて説明する。

【 0 2 3 8 】

先ず例えば、ユーザは、映像音声記録部 5 による映像信号及び音声信号を視聴することにより、検索したい CM を選択したとする。このとき、CM 特徴量比較器 1 1 1 には、ステップ S 1 2 0 として、上記ユーザによる選択に応じた検索指令が入力されることになる。

【 0 2 3 9 】

このとき、CM 特徴量比較器 1 1 1 は、ステップ S 1 2 1 の処理として、上記入力された検索指令に基づいて、CM データベース 1 1 0 から、その検索指令に該当する CM の特徴量 $G[n]$ を取得する。

【 0 2 4 0 】

続いて、CM特徴量比較器 1 1 1 は、ステップ S 1 2 2 として CM データベース 1 1 0 より、検索される候補 CM を一つ選択し、さらにステップ S 1 2 3 として、その候補 CM に対応する特徴量 $G'[n]$ を取得する。

【0 2 4 1】

次に、CM特徴量比較器 1 1 1 は、ステップ S 1 2 4 として、上記選択された CM について、式 (3 0) の計算を行い、それを予め定める所定の閾値 J_{thsd} と比較する。

【0 2 4 2】

【数 1 5】

$$J(G, G') = \sum_R \left| G[n] - G[n'] \right|^2 \quad (30)$$

【0 2 4 3】

ここで、このステップ S 1 2 4 において、 $J(G, G') < J_{thsd}$ と判定したならば (Y e s)、CM特徴量比較器 1 1 1 は、ステップ S 1 2 5 に進み、特徴量が一致したとして検索結果を出力し、ステップ S 1 2 2 に戻って再び他の候補 CM について同様の処理を行う。一方、ステップ S 1 2 4 において、 $J(G, G') < J_{thsd}$ でないと判定した場合 (N o)、CM特徴量比較器 1 1 1 は、特徴量が一致しなかったとして、ステップ S 1 2 2 に戻り、再び他の候補 CM について同様の処理を行う。

【0 2 4 4】

以上により、CM特徴量比較器 1 1 1 では、ユーザにより指定された CM と同じ CM を、映像音声記録部 5 に記録されているデータの中から検索することができる。

【0 2 4 5】

次に、図 2 5 には、上述した図 2 や図 1 5 に示した CM 検出部 4 を実装する場合のハードウェア構成の一例を示す。

【0 2 4 6】

この図 2 5 において、A/D変換器 4 0 は、前記図 2 や図 1 5 の A/D変換器 1 0 及び 1 3 の機能を備え、メモリ 4 1 は、前記フレームメモリ 1 1 及び音声信

号バッファ14の機能を備えている。

【0247】

A/VプロセッサまたはDSP（ディジタルシグナルプロセッサ）42は、前記カットチェンジ検出器112、振幅検出器15、相関検出器16、スペクトル検出器17、音源識別器101等の機能を備え、メモリ43は、前記特徴量バッファ18の機能を備えている。

【0248】

プロセッサ44は、前記CM候補検出器19、付加情報算出器20、付加条件判定器21、ルール判定器22、CM確率データベース103等の機能を備えている。

【0249】

前記動作制御部23の機能については、A/VプロセッサまたはDSP（ディジタルシグナルプロセッサ）42か、或いは、プロセッサ44が備えることができる。

【0250】

以上説明したような本発明の各実施の形態によれば、TV放送信号からCM部分を正確に検出可能とすることにより、例えばCMを不要としている視聴者や、CMを必要としている視聴者の双方に対して利便を図ることが可能となる。すなわち例えばCM部分を不要とする場合、テレビ放送信号からCM部分をスキップして視聴可能とする装置を実現でき、これは例えば番組本編のみを連続視聴する要求に対して有用な装置となる。また例えば、CM部分のみを必要とする場合、TV放送信号からCM部分のみを視聴できる装置を実現でき、これは例えばCMのみを連続視聴する要求に対して有用な装置となる。

【0251】

さらに、TV放送信号からCM部分を正確に検出可能とすることにより、例えば特定のCMの放送状況を調査する場合などにも有用となる。

【0252】

【発明の効果】

本発明の信号処理装置及び方法においては、入力信号から所定の時間間隔で発

生する信号の特徴的パターンに基づいて、第 1 の信号の候補区間を検出し、その候補区間内又はその前後の入力信号から第 1 の信号らしさを表わす特徴量を抽出し、その特徴量に基づき第 1 の信号の区間を検出するようにしたので、例えば、TV 放送信号に含まれるコマーシャルメッセージ部分を高精度に検出又は検索可能となる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施の形態の映像音声記録装置の概略構成図である。

【図 2】

第 1 の具体例の CM 検出部の詳細な構成図である。

【図 3】

CM 検出部のフロントエンド部における映像信号処理の流れを示すフローチャートである。

【図 4】

CM 検出部のフロントエンド部における音声信号処理の流れを示すフローチャートである。

【図 5】

CM 検出部の CM 候補検出器における動作の流れを示すフローチャートである。

【図 6】

必須条件の算出例の説明に用いる図である。

【図 7】

第 1 の具体例の CM 検出部における CM 候補テーブルを示す図である。

【図 8】

CM 検出部の付加条件算出器における特徴量の算出例の説明に用いる図である。

【図 9】

付加条件算出器の構成図である。

【図 10】

スコア算出演算の際の単位ステップ関数、矩形関数、シグモイド型関数の説明に用いる図である。

【図 1 1】

ルール判定器の動作の流れを示すフローチャートである。

【図 1 2】

最小長さ優先ルールの説明に用いる図である。

【図 1 3】

隣接優先ルールの説明に用いる図である。

【図 1 4】

スコア優先ルールの説明に用いる図である。

【図 1 5】

第 2 の具体例の CM 検出部の詳細な構成図である。

【図 1 6】

第 2 の具体例の CM 検出部における CM 候補テーブル（拡張部分のみ）を示す図である。

【図 1 7】

第 3 の具体例の CM 検出部における CM 候補テーブル（拡張部分のみ）を示す図である。

【図 1 8】

CM 検出部の付加条件算出器における小振幅回数計算処理を説明するフローチャートである。

【図 1 9】

小振幅回数の計算の具体例を示す図である。

【図 2 0】

CM 検出部の付加条件算出器における小振幅区間長計算処理を説明するフローチャートである。

【図 2 1】

小振幅区間長の具体例を示す図である。

【図 2 2】

振幅分散の具体例を示す図である。

【図 2 3】

本発明の第 2 の実施の形態の映像音声記録装置の概略構成図である。

【図 2 4】

CM 特徴量比較器の動作の流れを示すフローチャートである。

【図 2 5】

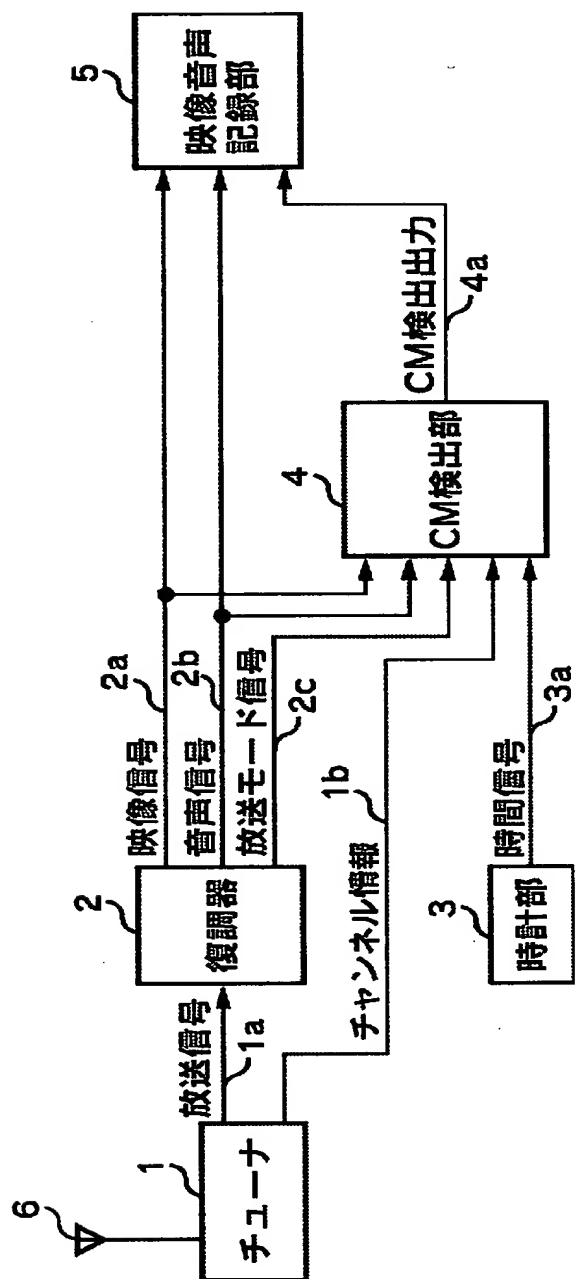
CM 検出部を実装する場合の一例としてのハードウェア構成図である。

【符号の説明】

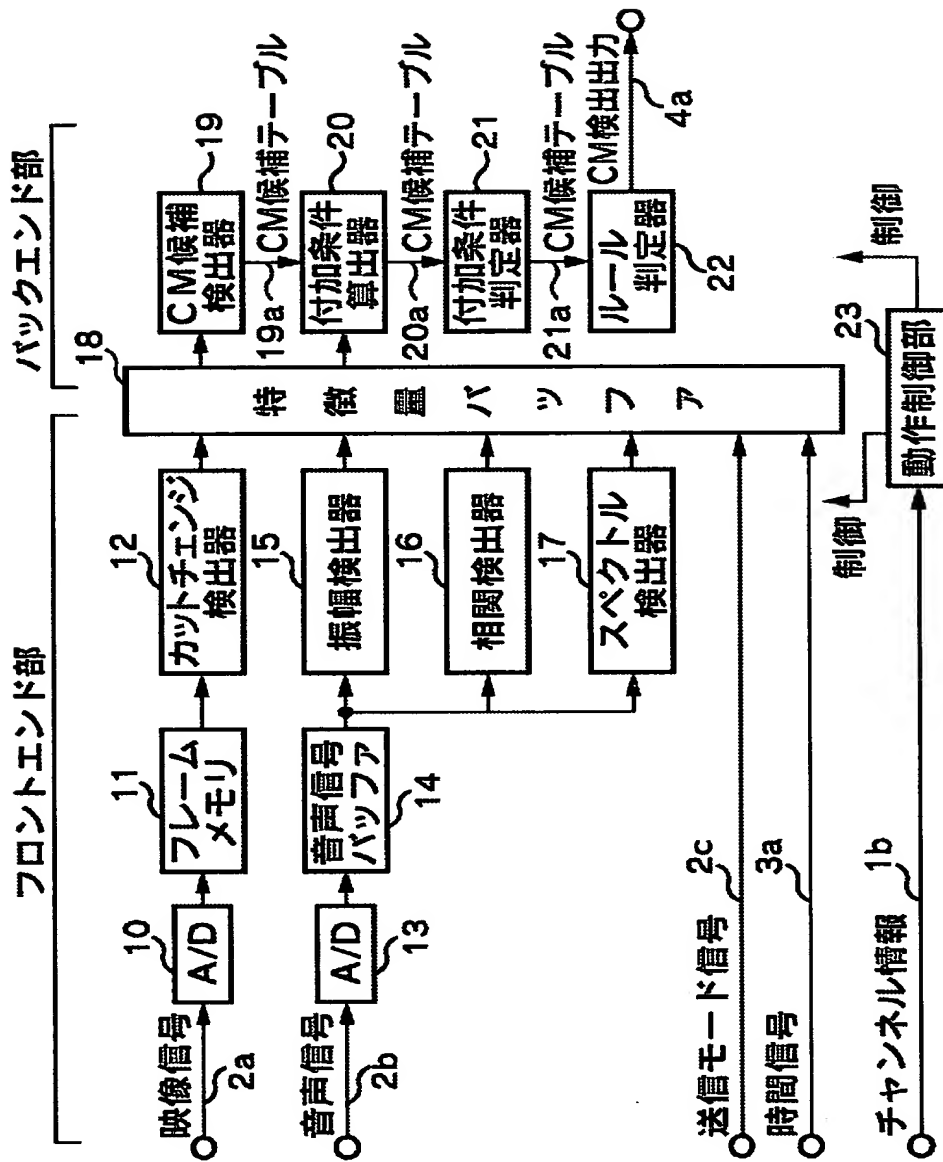
1 チューナ、 2 復調器、 3 時計部、 4 CM 検出器、 5 映像
音声記録部、 10, 13 A/D 変換器、 11 フレームメモリ、 14
音声信号バッファ、 12 カットチェンジ検出器、 15 振幅検出器、 1
6 相関検出器、 17 スペクトル検出器、 18 特徴量バッファ、 19
CM 候補検出器、 20 付加条件算出器、 21 付加条件判定器、 22
ルール判定器、 23 動作制御部、 101 音源識別器、 102 番組
ジャンルデータまたは番組ジャンル識別器、 103 CM 確率データベース、
110 CM データベース、 111 CM 特徴量比較器

【書類名】 図面

【図 1】

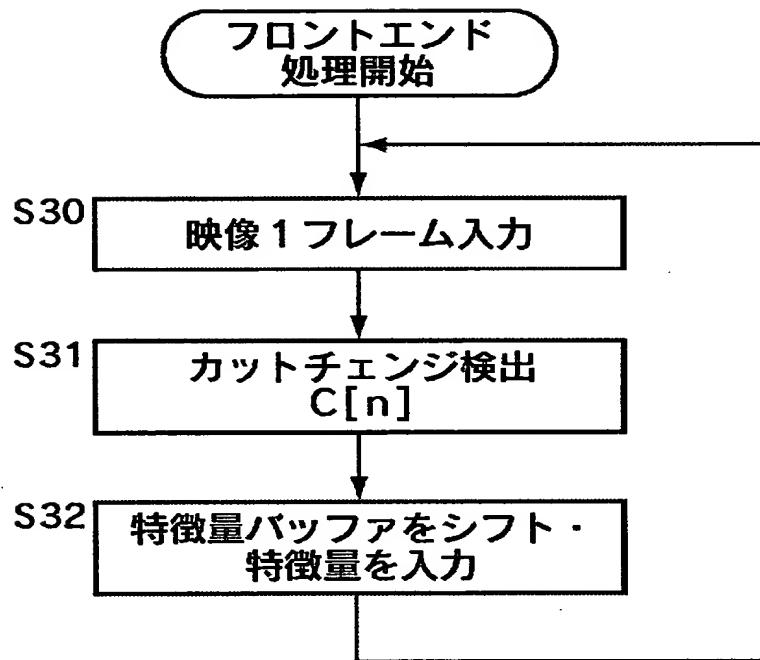


【図 2】

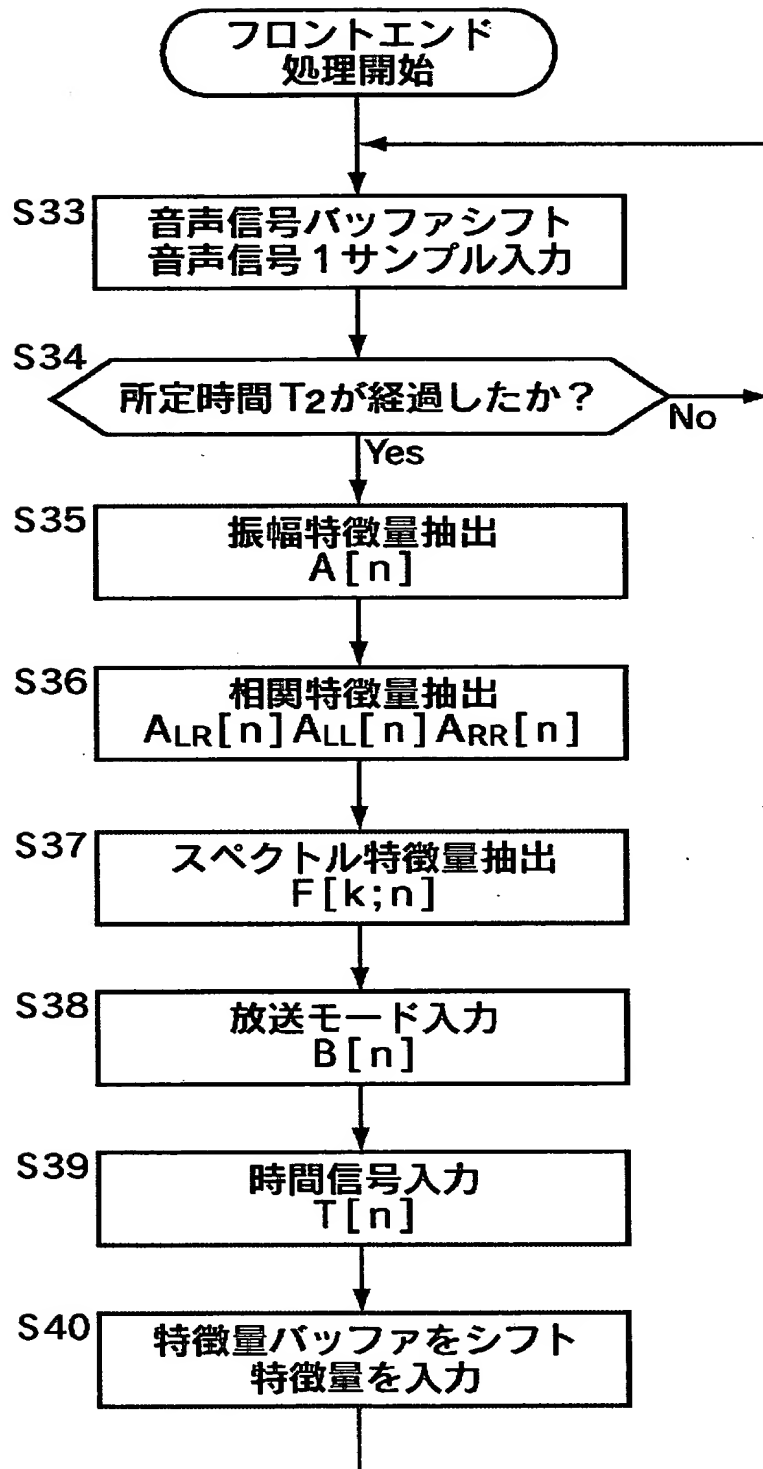


CM検出部 4

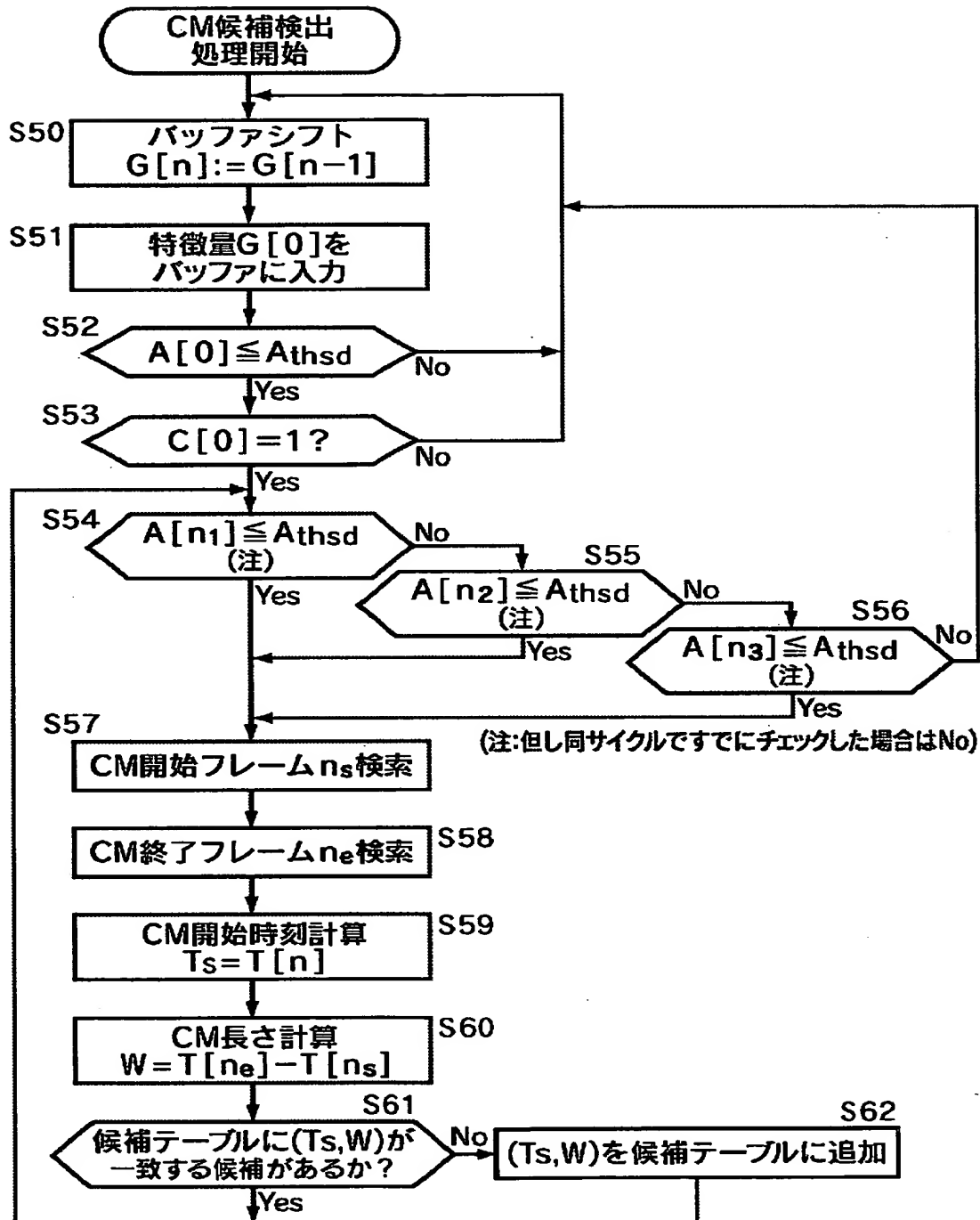
【図 3】



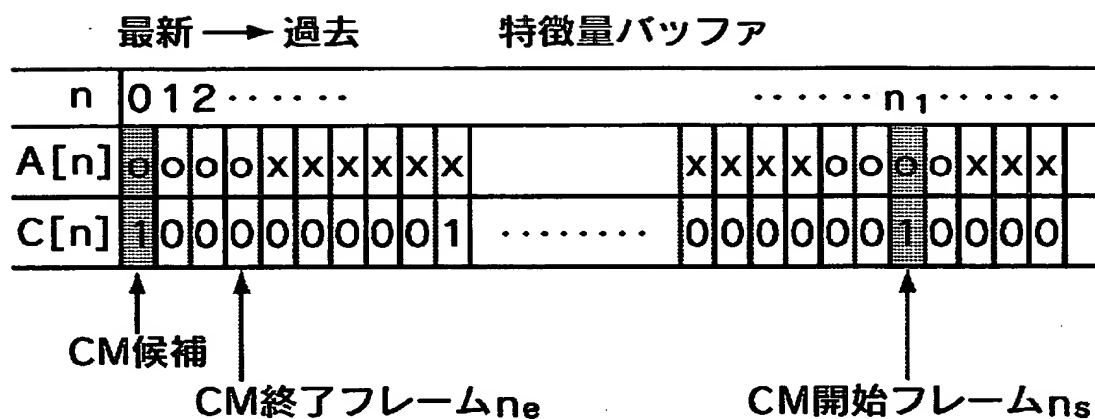
【図 4】



【図 5】



【图 6】

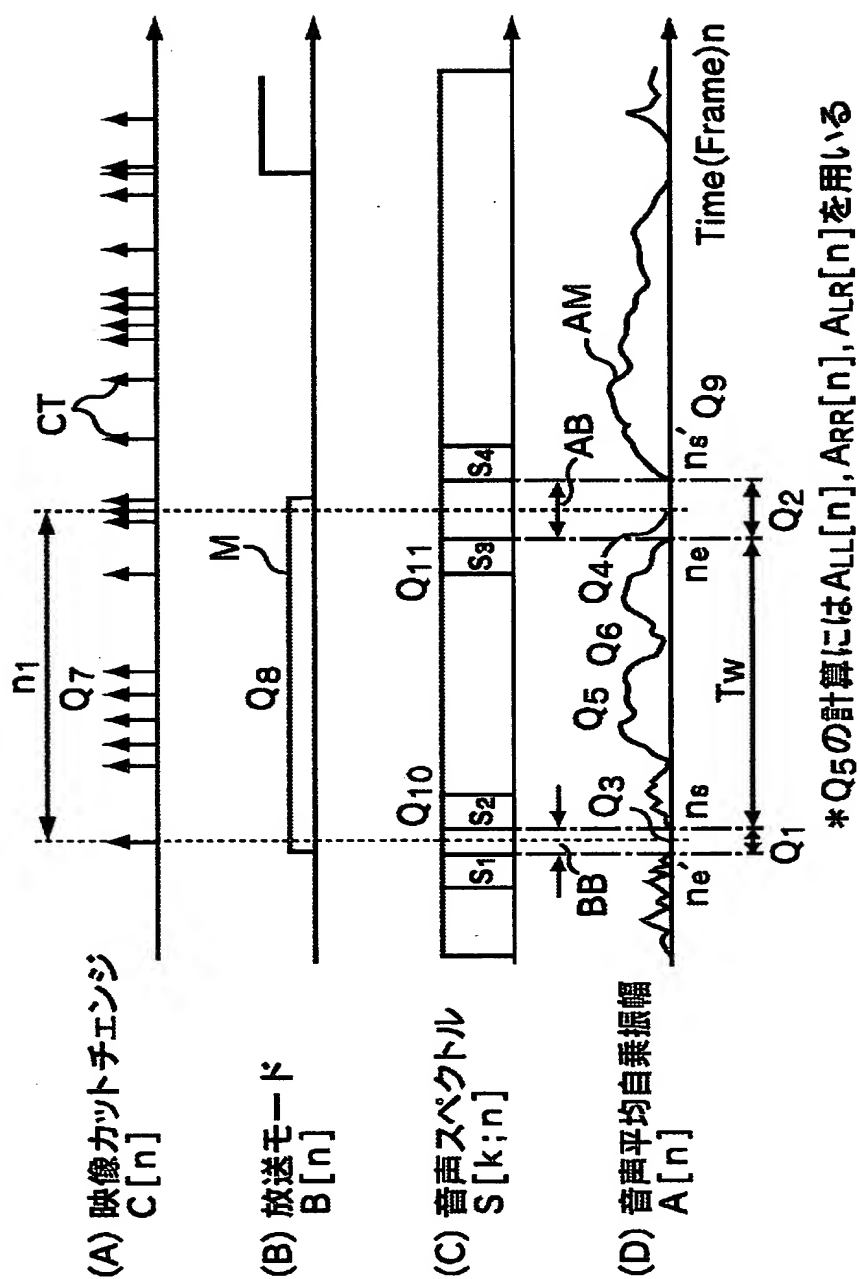


【图 7】

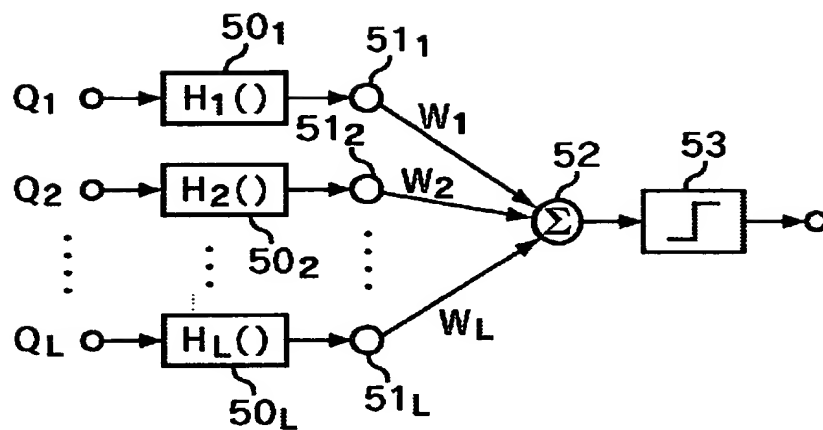
項 目	記号	単位	必須条件例 (19a)	必須条件例 (20a)	判定例 (21a)
開始時刻	Ts	時分秒	1:23'45	1:23'45	1:23'45
長さ(有音長)	Tw	秒	14.63	14.63	14.63
前ブレーク長	Q ₁	ms	—	300.0	300.0
後ブレーク長	Q ₂	ms	—	300.0	300.0
前ブレーク最小振幅	Q ₃	注	—	0.00015	0.00015
後ブレーク最小振幅	Q ₄	注	—	0.00020	0.00020
左右相関値	Q ₅	—	—	0.934	0.934
平均振幅値	Q ₆	注	—	0.010	0.010
カット数	Q ₇	個	—	9	9
放送モード	Q ₈	—	—	1	1
隣接候補数	Q ₉	個	—	2	2
前スペクトル差分エネルギー	Q ₁₀	—	—	0.41	0.41
後スペクトル差分エネルギー	Q ₁₁	—	—	0.63	0.63
スコア	R	—	—	—	1.80
スコア判定結果	Z	—	—	—	1

注：ここでの音声信号の振幅に関する量は、最大振幅に対する比率として表している

【図 8】

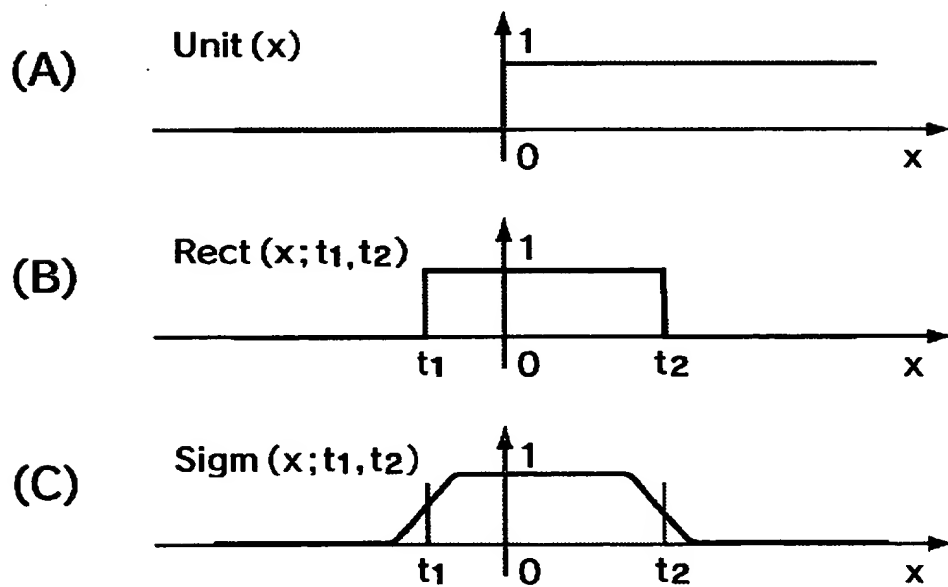


【図 9】

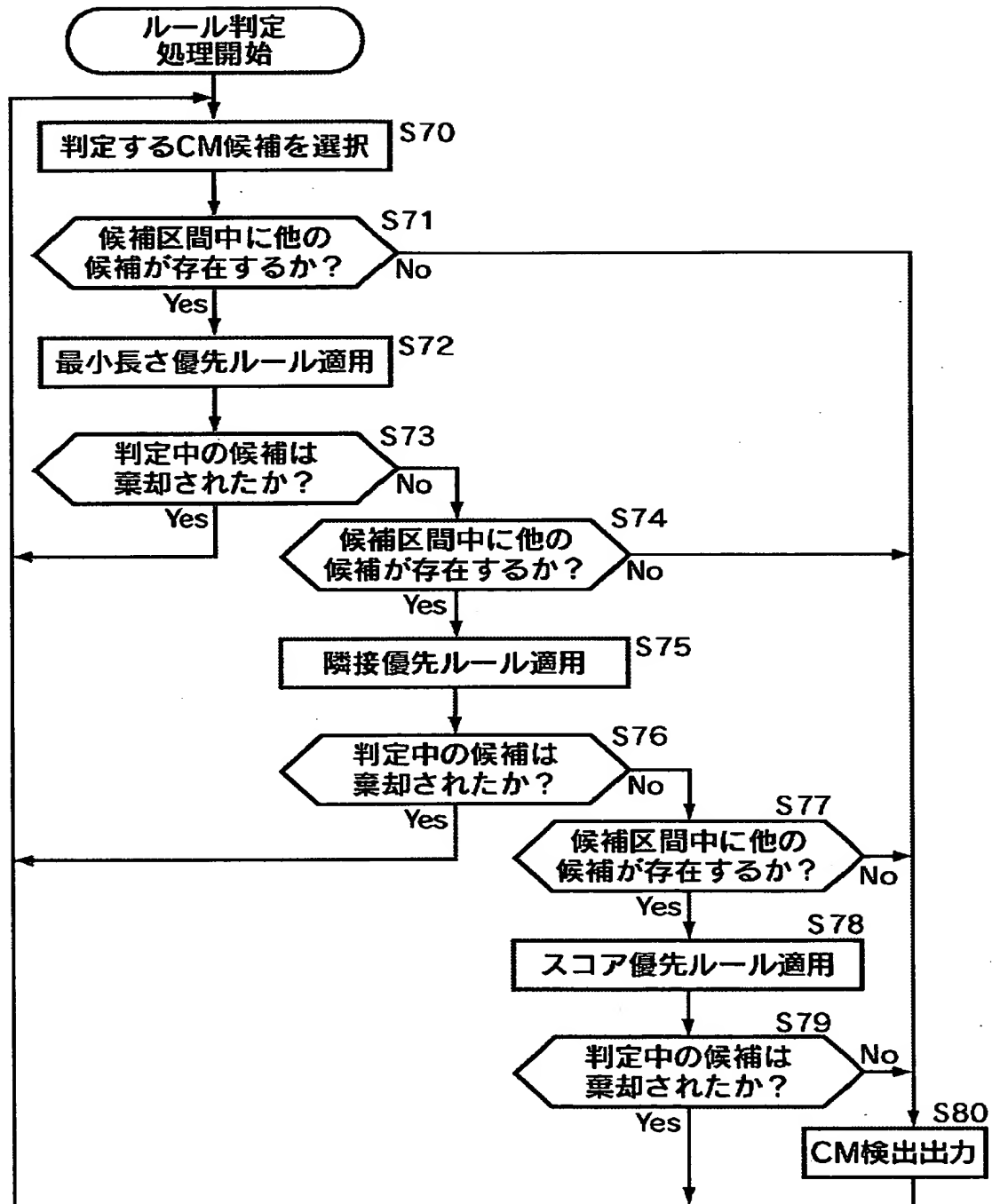


付加条件判定器 21

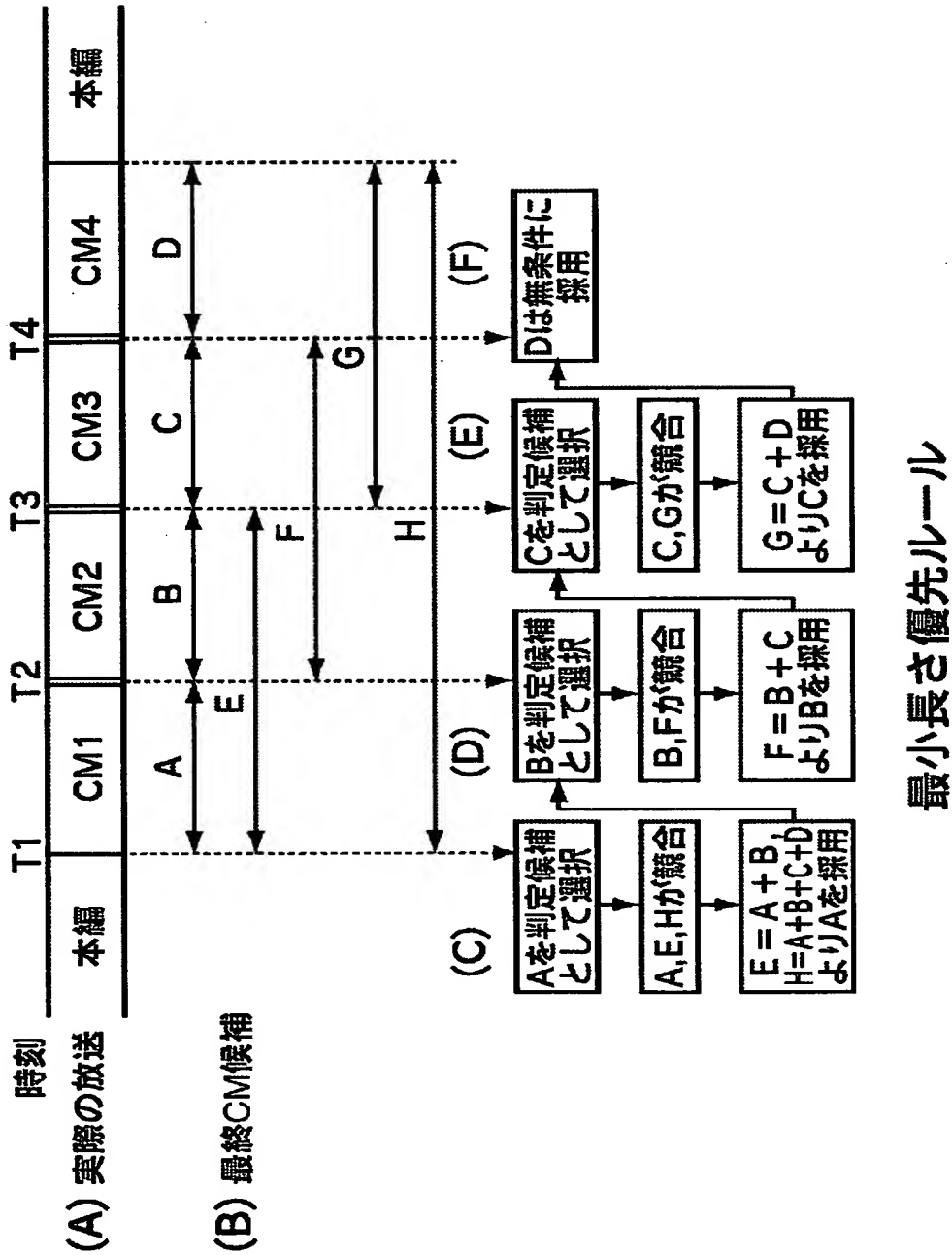
【図 1 0】



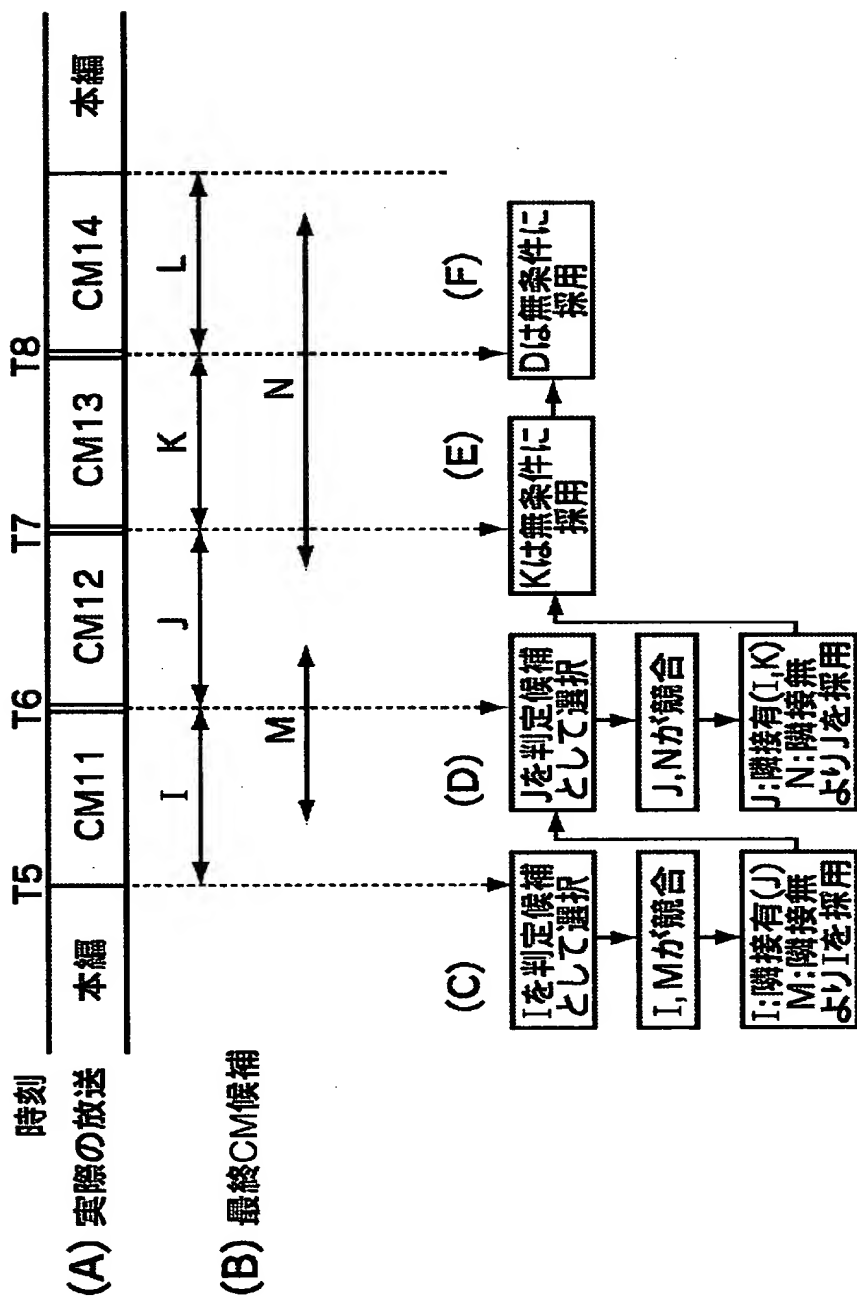
【図 11】



【图 1 2】

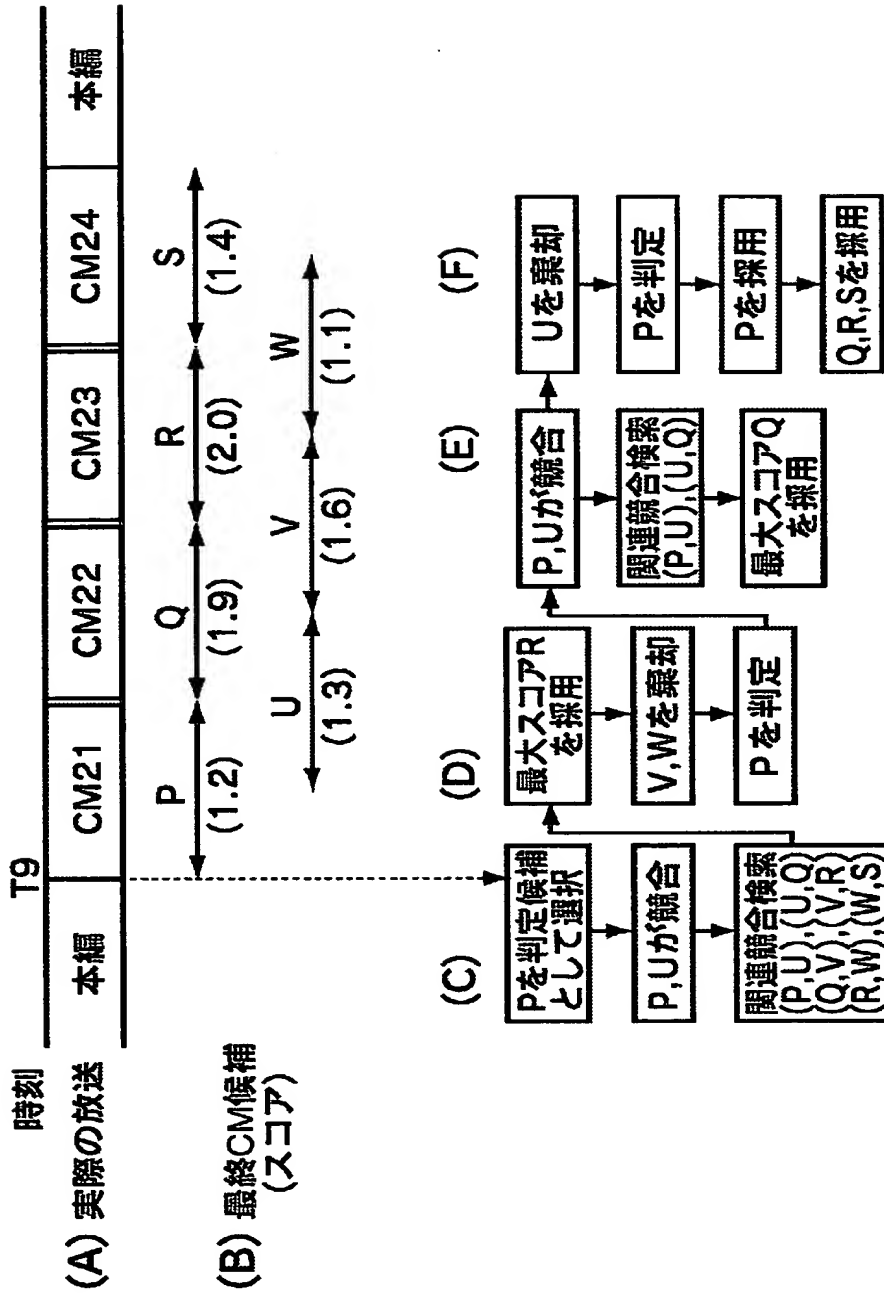


【図 1 3】



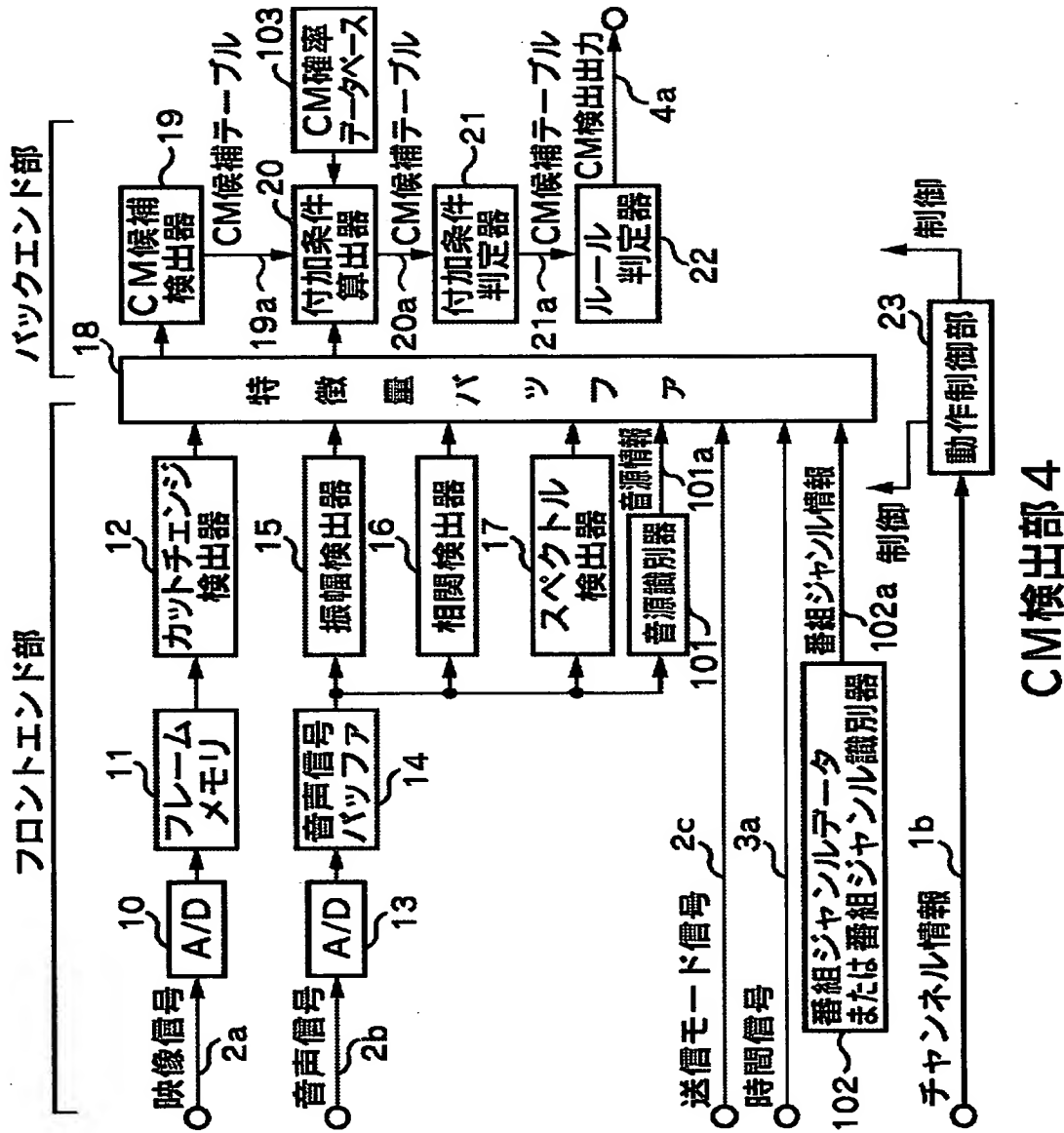
隣接優先ルール

【図14】



スコア優先ルール

【図15】



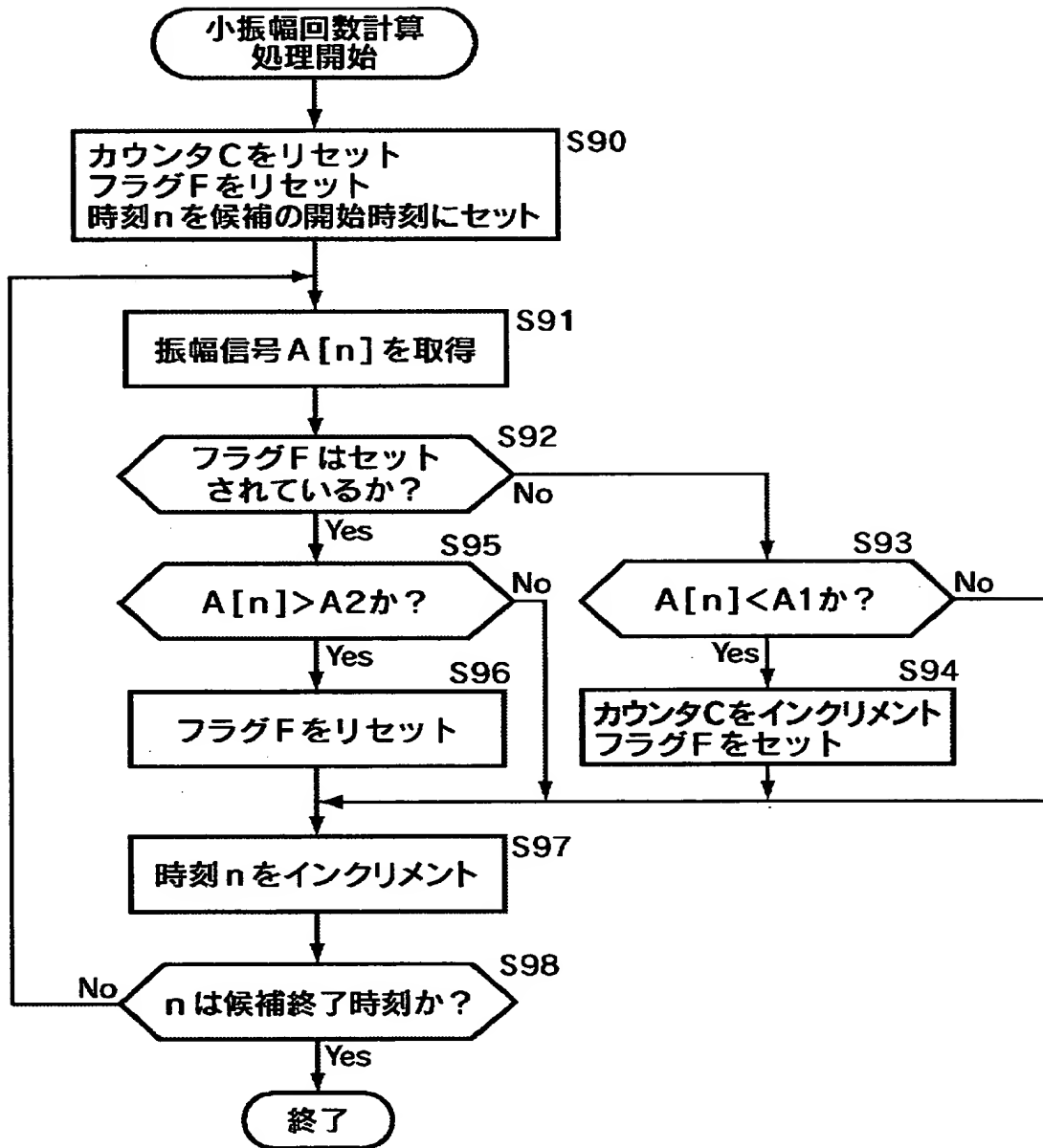
【図 16】

項 目	記号	単位	必須条件例 (19a)	付加条件例 (20a)	判定例 (21a)
音声の有無	Q12	—	—	1	1
音楽の有無	Q13	—	—	1	1
時間帯確率	Q14	—	—	0.15	0.15
番組ジャンル確率	Q15	—	—	0.1	0.1

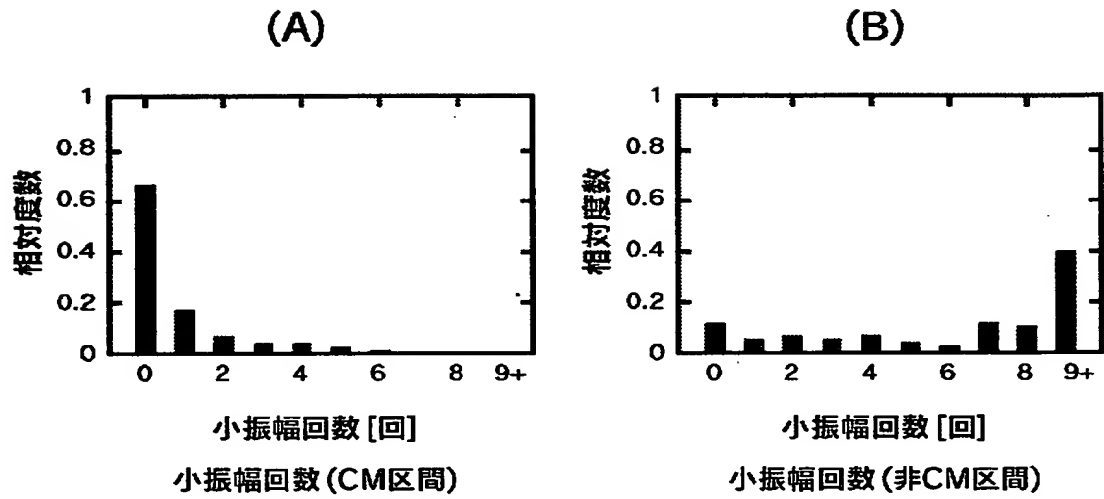
【図 17】

項 目	記号	単位	数値例
小振幅回数	Q16	—	1
小振幅区間	Q17	s	0.24
信号分散	Q18	—	0.40

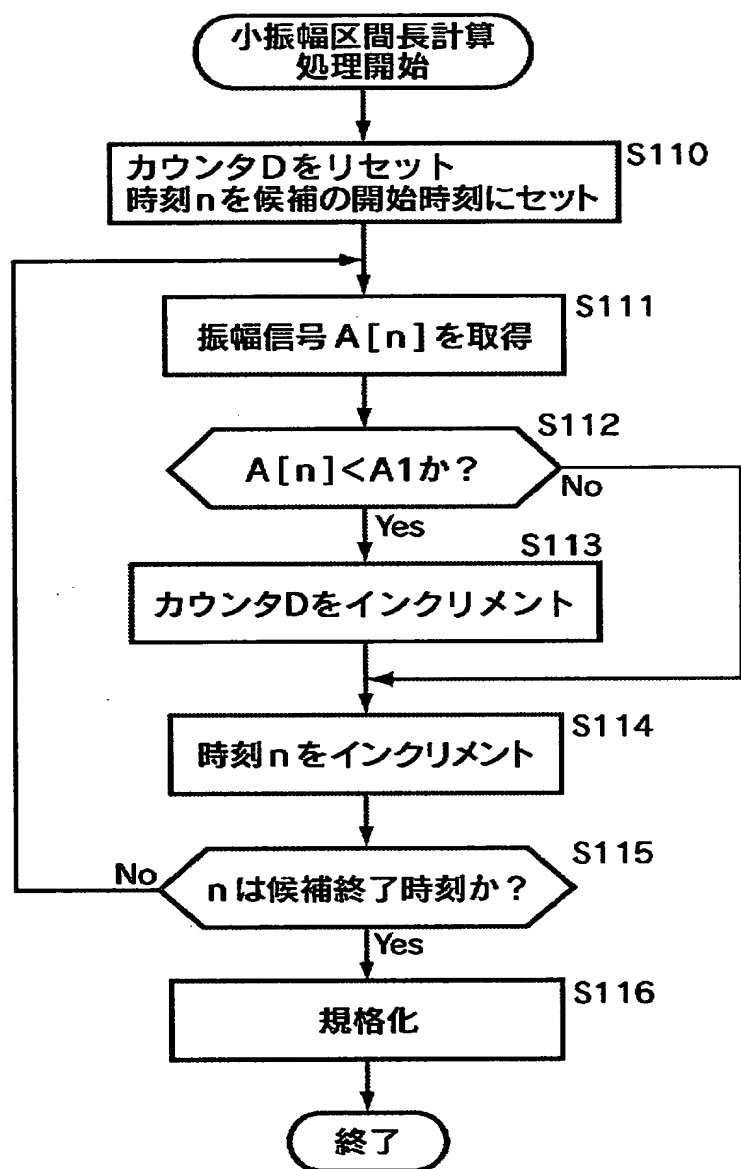
【図18】



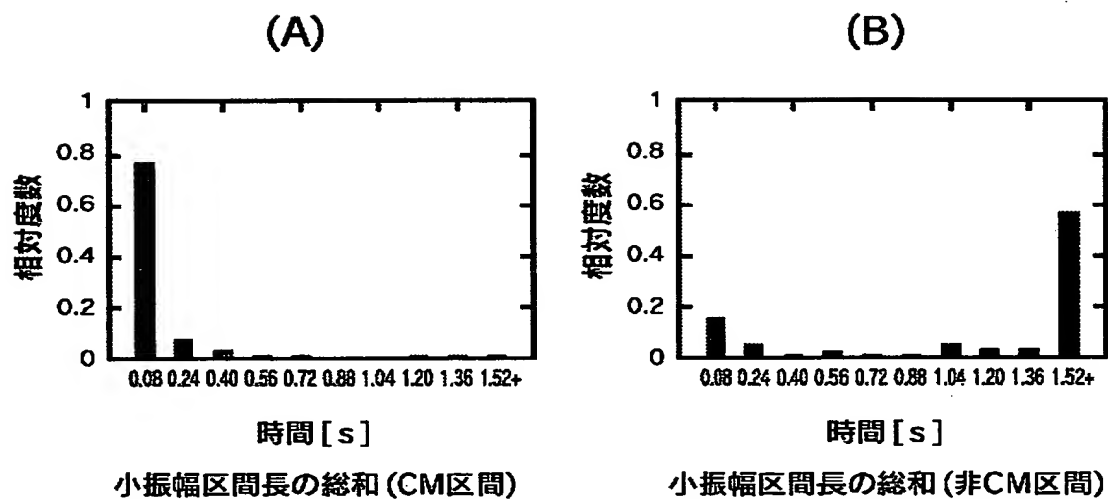
【図 1 9】



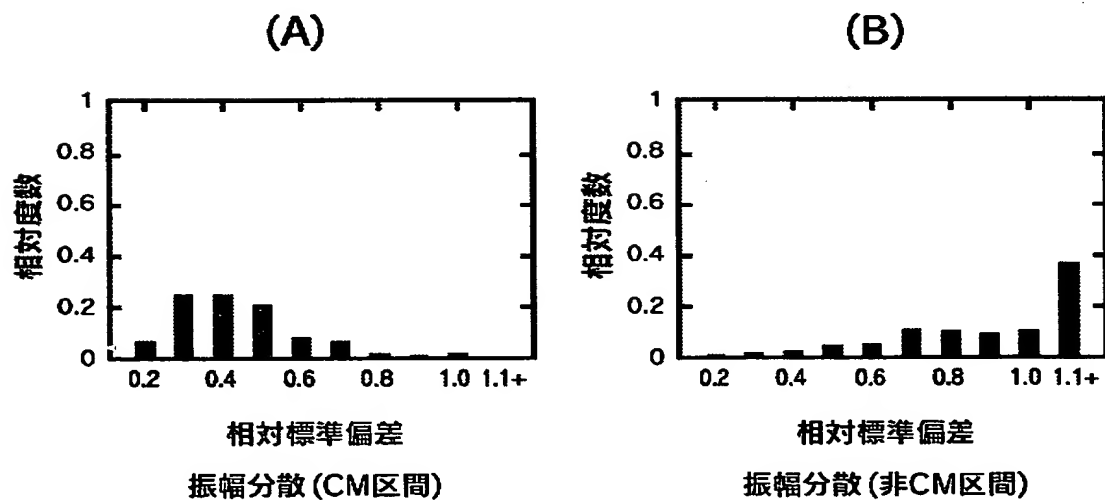
【図 2 0】



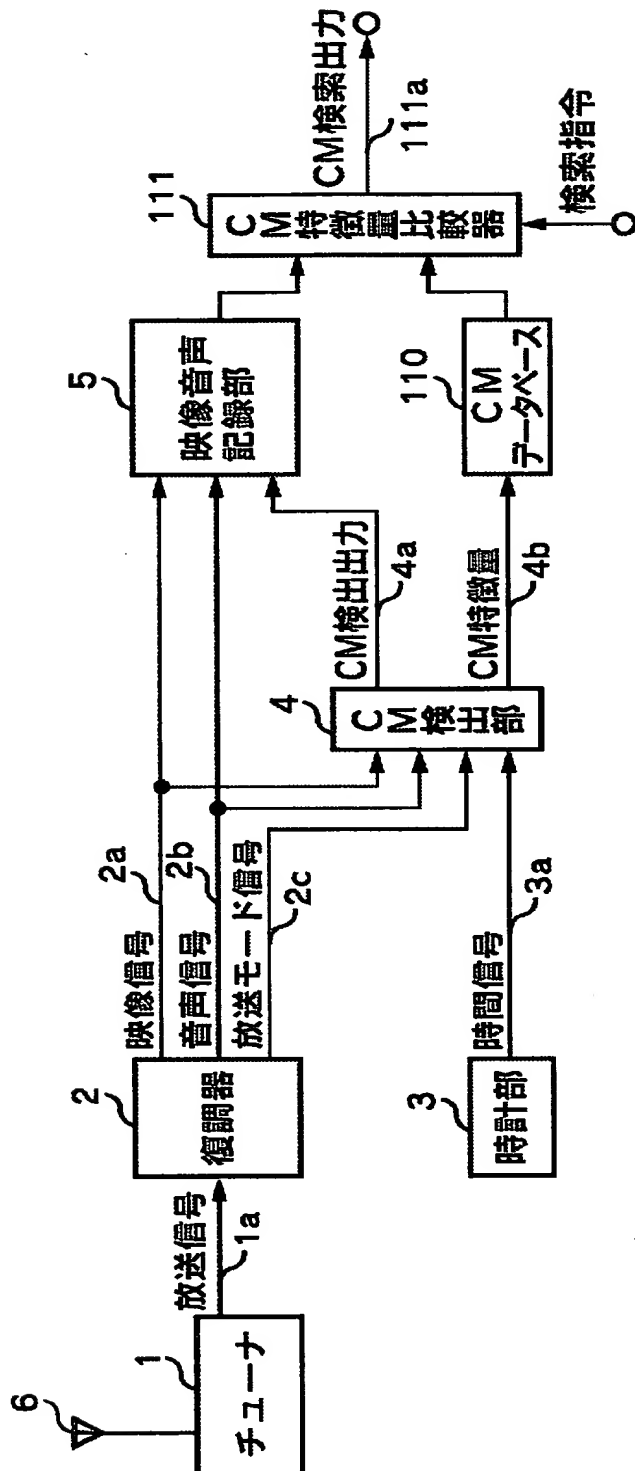
【図 2 1】



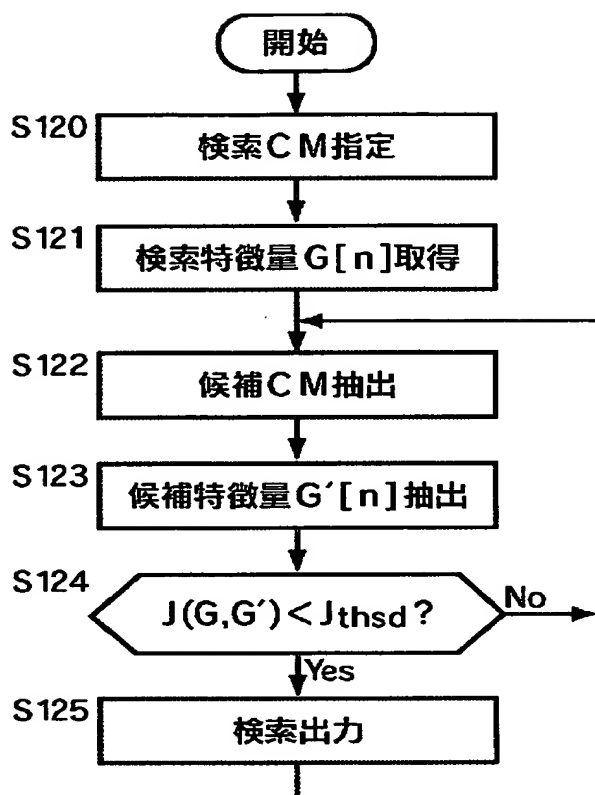
【図 2 2】



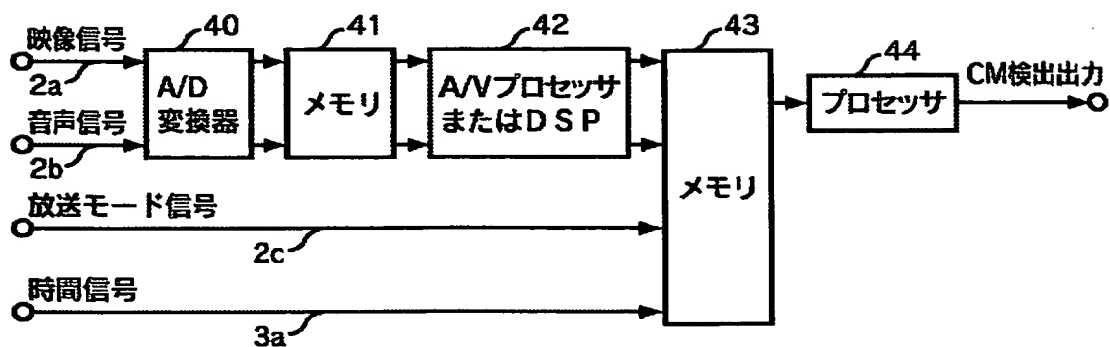
【図 23】



【図 2 4】



【図 2 5】



【書類名】 要約書

【要約】

【課題】 例えばＴＶ放送信号のＣＭ部分を高精度に検出又は検索可能とする。

【解決手段】 アンテナ６を介しチューナ１にて受信し、さらに復調器２にて復調されたＴＶ放送信号の音声及び映像信号から、１５，３０，６０秒間隔で発生する特徴的パターンに基づいて、ＣＭの候補区間を検出し、そのＣＭ候補区間内又はその前後の音声及び映像信号から、ＣＭらしさを表わす特徴量を抽出し、その特徴量に基づきＣＭ区間を検出するＣＭ検出部４を備える。

【選択図】 図１

認定・付加情報

特許出願の番号	特願 2000-307907
受付番号	50001301162
書類名	特許願
担当官	第七担当上席 0096
作成日	平成12年10月12日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000002185
【住所又は居所】	東京都品川区北品川6丁目7番35号
【氏名又は名称】	ソニー株式会社

【代理人】

申請人	
【識別番号】	100082131
【住所又は居所】	東京都新宿区西新宿7丁目5番8号 GOWA西 新宿ビル6F 稲本国際特許事務所
【氏名又は名称】	稲本 義雄

出 願 人 履 歴 情 報

識別番号 [000002185]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都品川区北品川6丁目7番35号
氏 名 ソニー株式会社